

FIGURE 1
Plasmid sequence of pNC5LSPCEAp53 (pMC30B5) for vCP2086

```

5      1  GCCCTTT CGTCTCG CGCGTTT CGGTGAT GACGGTG AAAACCT CTGACAC ATGCAGC TCCCGGA GACGGTC
      71  CGGGAAA GCAGAGC GCGCAAA GCCACTA CTGCCAC TTTTGGG GACTGTG TACGTGC AGGGGCT CTGCCAG
      141  ACAGCTT GTCTGTA AGCGGAT GCCGGGA GCAGACA AGCCCGT CAGGGCG CGTCAGC GGGTGTG GCGGGGT
      211  TGTCGAA CAGACAT TCGCCTA CGGCCCT CGTCTGT TCGGSCA GTCCCGC GCAGTCG CCCACAA CCGCCCA
      281  GTCGGGG CTGGCTT AACTATG CGGCATC AGAGCAG ATTTGAC TGAGAGT GCACCAT ATGCGGT GTGAAAT
      351  CAGCCCC GACCGAA TTGATAC GCCGTAG TCTCGTC TAACATG ACTCTCA CGTGGTA TACGCCA CACTTTA
      421  ACCGCAC AGATGCG TAAGGAG AAAATAC GGCATCA GCGGCCA TTGCGCA TTCAGGC TCGGCAA CTGTTGG
      491  TGGCGTG TCTACGC ATTCCTC TTTTATG GCGTAGT CCGCGGT AAGCGGT AAGTCCG ACGCGTT GACAACC
      561  GAAGGGC GATCGGT GCGGGCC TCTTCGC TATTACG CCAGCTG GCGAAGG GGGGATG TGCTGCA AGGCGAT
      631  CTTCOCG CTAGCCA CGCCCGG AGAAGCG ATAATGC GGTGCGC CGCTTTC CCCCTAC ACGACGT TCCGCTA
      701  TAAGTTG GGTAAAG CCAGGGT TTCCCA GTCACGA CGTTGTA AAACGAC GGCCAGT GCCAAGC TTGGCTG
      771  ATTCAAC CCATTGC GGTCCCA AAAGGGT CAGTGCT GCAACAT TTTGCTG COGGTCA CGGTTCC AACCGAC

      ~~~~~~
      841  CAGGTAT TCTAAAC TAGGAAT AGATGAA ATTATGT GCAAAGG AGATAAC TTTAGAT ATGGATC TGAATTTA
      911  GTCCATA AGATTTG ATCCTTA TCTACTT TAATACA CGTTTCC TCTATGG AAATCTA TACCTAG ACTAAAT

      ~~~~~~
      981  TTTGGTT TTTTATA ATCATAA TCTAACA ACATTTT CACTATA CTATACC TTCCTGC ACAAGTC GCCATTA
      1051  AAACCAA AAGTAT TAGTATT AGATTGT TGTAAAA GTGATAT GATATGG AAGAACG TGTTGAG CGGTAAT

      ~~~~~~
      1121  GTAGIAT AGACTTA TACTTTG TAACCAT AGTATAC TTTAGCG CGTCATC TTCCTCA TCTAAAA CAGATTT
      1191  CATCATA TCTGAAT ATGAAAC ATTGGTA TCATATG AAATGCG GCAGTAG AAGAAGT AGATTTT GTCTAAA

      ~~~~~~
      1261  ACAACAA TAATCAT CGTCGTC ATCTTCA TCTTCAT TAAAGTT TTCAATAT TCAATAA CTTTCIT TTCTAAA
      1331  TGTGTTT ATTAGTA GCAGCAG TAGAAGT AGAAGTA ATTTCAA AAGTATA AGTTATT GAAAGAA AAGATTT

      ~~~~~~
      1401  ACATCAT CTGAATC AATAAAC ATAGAAC GGTATAG AGCGTTA ATCTCCA TTGTAAA ATATACT AACCGGT
      1471  TGTAGTA GACTTAG TTATTTG TATCTTG CCATATC TCGCAAT TAGAGGT AACATTT TATATGA TTGCGCA

      ~~~~~~
      1541  TGCTCAT GATGTAC TTTTITT CATTATT TAGAAAT TATGCAT TTTAGAT CTTTATA AGCGGCC GTGATTA
      1611  ACGAGTA CTACATG AAAAAA GTAATAA ATCTTTA ATACGTA AAATCTA GAAATAT TCGCCGG CACTAAT

      ~~~~~~
      1681  ACTAGTC ATAAAAA CCGGGGA TCGATTG TAGACTC GAGATAA AAACATAT ATCAGAG CAACCCC AACCGAG
      1751  TGATCAG TATTTTT GGGCCCT AGCTAAG ATCTGAG CTCIATT TTTGATA TAGTCTC GTTGGGG TTGGTGC

      ~~~~~~
      1821  ACTCAA TCATGAT GCGGACA GTGGCCC CAGCTGA GAGACCA GGAGAAG TTCCAGA TGCAAGG ACTGTGA
      1891  TGAGGTT AGTACTA CGGCTGT CACCGGG GTCGACT CTCTGCT CCTCTTC AAGGTCT ACGTCTC TGACACT

      ~~~~~~
      1961  ..GlyIle MetIle GlyValThr AlaGly AlaSer LeuGlyPro SerThr GlySer AlaSerVal ThrIle
      2031  TGCTCTT GACTATG GAATAT TCGCGCC AGTAGCC AAGITAG AGACAAA ACAGGCA TAGGTCC CGTTATT
      2101  ACGAGAA CTGATAC CTTAATA ACGCCGG TCATCGG TTCAATC TCTGTTT TGTCCGT ATCCAGG GCAATAA

      ~~~~~~
      2171  ..SerLys ValIleSer AsnAsn ArgGly ThrAlaLeu AsnSer ValPhe CysAlaTyr ThrGly AsnAsn
      2241  ATTTGGC GTGATT TGGCGAT AAAGAGA ACTGTG TGTGTTG CTGCGGT ATCCCAT TGATACG CCAAGAA
      2311  TAAACCG CACTAAA ACCGCTA TTTCTCT TGAACAC ACACAAC GACGCCA TAGGGTA ACTATGC GGTCTCT

      ~~~~~~
      2381  AsnProThr IleLys AlaIle PheLeuVal GlnThr HisGln GlnProIle GlyAsn IleArg TrpSerTyr
      2451  TACTGCG GGGATGG GTTAGAG GCCGAGT GGCAGGA GAGGTTG AGGTCCG CTCCCGA AAGGTAA GACGAGT
      2521  ATGACGC CCCTACC CAATCTC CGGCTCA CGTCCTT CTCCAAC TCCAGGC GAGGGCT TTCCATT CTGCTCA

      ~~~~~~
      2591  ..GlnPro SerPro AsnSerAla SerHis CysSer LeuAsnLeu AspAla GlySer LeuTyrSer SerAsp
      2661  CTGGGGG GGAATG ATGGGGG TGTCCGG CCCATAG AGGACAT CCAGGGT GACTGGG TCACTGC GGTTCGC
      2731  GACCCCC CTTTTAC TACCCCC ACAGGCC GGGTATC TCCTGTA GGTCCCA CTGACCC AGTGACG CCAACCG

      ~~~~~~
      2801  .ProPro SerIleIle ProThr AspPro GlyTyrLeu ValAsp LeuThr ValProAsp SerArg AsnAla
      2871  ACTCACT GAGTTCT GGATTCC ACATACA TAGGCTC TTGCGTC ATTTCIT GTGACAT TGAATAG AGTGAGG
      2941  TGAGTGA CTCAGA CTAAGG TGTATG ATCCGAG AACGCG TAAAGAA CACTGTA ACTTATC TCATCC

      ~~~~~~
      3011  SerValSer AsnGln IleGly CysValTyr AlaArg AlaAsp AsnArgThr ValAsn PheLeu ThrLeuThr
      3081  GTCCTGT TGCCATT GGACAGC TGCAGCG TGGGACT GACTGGG AGGCTCT GACCATT TACCCAC CACAGGT
      3151  CAGGACA ACGGTAA CCTGTC ACGTCGG ACCCTGA CTGACCC TCCGAGA CTGGTAA ATGGGTG GTGTCCA

      ~~~~~~
      3221  ..ArgAsn GlyAsn SerLeuGln LeuArg ProSer ValProLeu SerGln GlyAsn ValTrpTrp LeuTyr
      3291  AGGTGTG GTTCTGA GCCTCAG GTTCACA GGTGAAG GCCACAG CATCCTT GTCCTCC ACGGGTT TGGAGTT
      3361  TCCACAA CAGACT CGGAGTC CAAGTGT CCACCTC CGGTGTC GTAGGAA CAGGAGG TGCCCAA ACCTCAA

      ~~~~~~
      3431  .ThrThr AsnGlnAla GluPro GluCys ThrPheAla ValAla AspLys AspGluVal ProLys SerAsn

```

1471	GTGTGCTG GAGATGG AGGGGCTT GGGCAGC TCCGCGG AACACGT TATTGTT TTAAC TG TAGTCT GCTGTGA CAACGAC CTCTACC TCCCGAA CCGCTCG AGGCGCC TTGTGCA ATAACAA AATTGAC ATCAGGA CGACACT CEA
5	1541 AsnSerSer IleSer ProLys ProLeuGlu AlaSer ValThr IleThrLys ValThr ThrArg SerHisGly- CCACTGG CTGAGTT ATTGGCC TGGCAAG TATAGAG TCCGCTG TTCTTCT CAGTTAT GTTGCTT ATAAATA GGTGACC GACTCAA TAACCGG ACCGTTT ATATCTC AGGCGAC AAGAAGA GTCAATA CAACGAA TATTIAT CEA
10	1611 ..SerAla SerAsn AsnAlaGln CysThr TyrLeu GlySerAsn LysGlu ThrIle AsnSerIle PheLeu- ACTCTTG AGTATGC TGCTGAA TGTTTCC ATCAATC AGCCAGG AGTACTG TGCAGGG GGGTTGG ATGCTGC TGAGAAC TCATACG ACGACTT ACAAGG TAGTTAG TCGGTCC TCATGAC ACGTCCC CCCAACC TACGACG CEA
15	1681 .GluGln ThrHisGln GlnIle AsnGly AspIleLeu TrpSer TyrGln AlaProPro AsnSer AlaAla ATGGCAA GAAAGGC TCAAGTT CACGCCG GGACGGT AGTAGGT GTATGAT GGAGATA TAGTTGG GTCGTCT TACCGTT CTTTCCG AGTTCAA GTGCGGC CCTGCCA TCATCCA CATACTA CCTCTAT ATCAACC CAGCAGA CEA
20	1751 HisCysSer LeuSer LeuAsn ValGlyPro ArgTyr TyrThr TyrSerPro SerIle ThrPro AspAspPro- GGGCCAT ACAAAAC ATTAAGG ATAACAG GGTCGGA GTGATCA ACGGATA ATTCAAT CTGAATG CCACACT CCCGGTA TGTTTGG TAATTCC TATTGTC CCAGCCT CACTAGT TGCCAT TAAGTAA GACTTAC GGTGTGA CEA
25	1821 ..GlyTyr LeuVal AsnLeuIle ValPro AspSer HisAspVal SerLeu GluAsn GlnIleGly CysGlu- CATAAGG TCCTACA TCATTGC GAGTAAC GGACAGG AGTGTC AATGTGC GTTATCA TTAGACA ACTGCAA GTATTCC AGGATGT AGTAACG CTCATTG CCTGTCC TCACAGT TACACGC CAATAGT AATCTGT TGACGTT CEA
30	1891 .TyrPro GlyValAsp AsnArg ThrVal SerLeuLeu ThrLeu ThrArg AsnAspAsn SerLeu GlnLeu GCGTGGG CTAAACG GCAAACT TTGGTTA TTGACCC ACCATAA ATAAGTG GTATTTT GAATCTC TGGCTCA CGCACCC GATTGGC CGTTTGA AACCAAT AACTGGG TGGTATT TATTAC CATAAAA CTTAGAG ACCGAGT CEA
35	1961 ArgProSer ValPro LeuSer GlnAsnAsn ValTrp TrpLeu TyrThrThr AsnGln IleGlu ProGluCys- CAAGTTA ATGCAAC TCGGTCC TCATCCT CACTGG GTTAGAA TTGTAT TAGTTAT GAATGGT TTGGTGG GTTCAAT TACGTTG ACGCAGG AGTAGGA GTTGACC CAATCTT AACAAATG ATCAATA CTTACCA AAACCAC CEA
40	2031 ..ThrLeu AlaVal AlaAspGlu AspGlu ValPro AsnSerAsn AsnSer ThrIle PheProLys ProPro- GCTCATA CACGGA ATCGTC TCAAGGT TGTGCGG TTGAGTC CGGTGTC GCTATTG TGAGCTT GGCACGT CGAGTAT GTGCCAT TAGCAGC AGTGCCA ACACGCC AACTCAG GCCACAG CGATAAC ACTCGAA CCGTGCA CEA
45	2101 .GluTyr ValThrIle ThrThr ValThr ThrArgAsn LeuGly ThrAsp SerAsnHis AlaGln CysThr GTAGGAT CCACAT TGTTCAC GGTAAATA TTGGGAA TGAACAG TTCCTGG GTGGACT GTTGAA AGTGCCA CATCCTA GGTGATA ACAAGTG CCATTAT AACCCCT ACTGTG AAGGACC CACTGA CAACCTT TCACGCT CEA
50	2171 TyrSerGly SerAsn AsnVal ThrIleAsn ProIle PheLeu GluGlnThr SerGln GlnPhe ThrGlyAsn- TTGACAA ACCAGCT GTATTGG GCGGGAG GATTGCT AGCGGCA TGACAGC TCAGATT CAGATTT TCCCTTG AACTGTT TGCTGA CATAACC GCGCCCT CTAACGA TCGCCGT ACTGTGC AGTCTAA GTCTAAA AGGGGAC CEA
55	2241 ..ValPhe TrpSer TyrGlnAla ProPro AsnSer AlaAlaHis CysSer LeuAsn LeuAsnGlu GlySer- ATCTATA GCTTGTG TTTAGAG GGCTGAT TGTAGGA GCATCGG GTCCGTA AAGCAGG TTGAGAA TCACTGA TAGATAT CGAACAC AAATCTC CCGACTA ACATCCT CGTAGCC CAGGCAT TTCGTGC AACTCTT AGTGACT CEA
60	2311 .ArgTyr SerThrAsn LeuPro SerIle ThrProAla AspPro GlyTyr LeuValAsn LeuIle ValSer ATCAGAC CTCCTGG CGCTGAC TGGATTT TGGGTTT CGCATTT GTAGCTT GCTGTGT CGTTCTT GGTACG TAGTCTG GAGGACC GCGACTG ACCTAAA ACCCAA GCGTAAA CATCGAA CGACACA GCAAGGA CCGTGC CEA
65	2381 AspSerArg ArgAla SerVal ProAsnGln ThrGlu CysLys TyrSerAla ThrAsp AsnArg ThrValAsn- TTAAACA GGGTCAG AGTTCTA TTCCGT TGCTGAG TTGGAGT CTAGGGG ACACAGG CAGGGAC TGTTGT AAITTTGT CCCAGTC TCAAGAT AAAGGCA ACGACTC AACCTCA GATCCCC TGTGTCC GTCCCTG ACCAACA CEA
70	2451 ..PheLeu ThrLeu ThrArgAsn GlyAsn SerLeu GlnLeuArg ProSer ValPro LeuSerGln AsnAsn- TCACCCA CCAGAGA TATGTTG CGTCTTG AGTTTCG GGCTCGC ATGTAAA AGCGAGC GCATCTT TGTCTTC AGTGGGT GGTCTCT ATACAAC GCAGAAC TCRAAGC CCGAGCG TACATTT TCGCTGC CGTAGAA ACAGAG CEA
75	2521 .ValTrp TrpLeuTyr ThrAla AspGln ThrGluPro GluCys ThrPhe AlaValAla AspLys AspGlu GACAGGC TTACTAT TATTGGA GCTAATA GAAGGCT TAGGGAG TTCGGG TATACCC GGAAGCT GCCAGTT CTGTCCG AATGATA ATAACCT CGATTAT CTTCGGA ATCCCTC AAGGCCC ATATGGG CCTTGAC CGTCAA CEA
80	2591 ValProLys SerAsn AsnSer SerIleSer ProLys ProLeu GluProTyr ValArg PheGln GlyThrAla- GCTTCTT CATTAC AGATCT GACTTTA TGACGTG TAGGGTG TAGAATC CTGTGTC ATTCTGG ATGATGT CGAAGAA GTAAGTG TTCTAGA CTGAAGT ACTGCAC ATCCAC ATCTTAG GACACAG TAAGACC TACTACA CEA
85	2661 ..GluGlu AsnVal LeuAspSer LysIle ValHis LeuThrTyr PheGly ThrAsp AsnGlnIle IleAsn- TCTGGAT CAGCAGG GATGCAT TGGGTA TATTATC TCTGAC CACTGTA TCGGGG CCTGGG TAGCTG AGACCTA GTCGTCC CTACGTA ACCCAT ATAATAG AGAGCTG GTGACAT ACGCCC GGACCCC ATCGAAC CEA
90	2731 .GlnIle LeuLeuSer AlaAsn ProTyr IleIleGlu ArgGly SerTyr AlaProGly ProThr AlaGln TTAGATT CCTATTA CATATCC TATAATT TGACGGT TGCCATC CACTCTT TCACCTT TGTACCA GCTGTAG AACTCAA GGATAAT GTATAGG ATATTAA ACTGCCA ACGGTAG GTGAGAA AGTGGAA ACATGGT CGACATC

CEA
2801 GlnThrGly IleVal TyrGly IleIleGln ArgAsn GlyAsp ValArgGlu GlyLys TyrTrp SerTyrGly
CCAAAAA GATGCTG GGGCAGA TTGTGGA CAAGTAG AAGCACC TCCTTCC CCTCTGC GACATTG AACGGCG
GGTTTTT CTACGAC CCGTCT AACACCT GTTCATC TCGTGG AGGAAGG GGAGACG CTGTAAC TTGCCGC

5
CEA
2871 ..PheLeu HisGln ProLeuAsn HisVal LeuLeu LeuValGlu LysGly GluAla ValAsnPhe ProThr
TGGATTG AATAGTG AGCTTGG CAGTGGT GGGCGGG TTCCAGA AGGTTAG AAGTGAG GCTGTGA GCAGGAG
ACCTAAG TTATCAC TCGAACC GTCACCA CCGCGCC AAGGTCT TCCAATC TTCCTC GCACACT CGTCTC

10
CEA
2941 .SerGlu IleThrLeu LysAla ThrThr ProProAsn TrpPhe ThrLeu LeuSerAla ThrLeu LeuLeu
CCTCTGC CAGGGGA TGCACCA TCTGTGG GGAGGGG CCGAGGG AGACTCC ATTATTT ATATTCC AAAAAA
GGAGACG GTCCCCC ACGTGGT AGACACC CCTCCCC GGCTCCC TCTGAGG TAATAAA TATAAGG TTTTTT

15
E/L Promoter
CEA
ArgGlnTrp ProIle CysTrp ArgHisPro ProAla SerPro SerGluMet
H6 promoter
20 3011 AAAAAA AAATTTC AATTTT GTCGACC TGCAGCT CGACGGA TCCCCC GGGTCT TTATTCT ATACTTA
TTTTAT TTTAAG TTAATAA CAGCTGG ACGTGA GCTGCCT AGGGGGG CCAAGA AATAAGA TATGAAT

25
E/L Promoter
H6 promoter
3081 AAAAGTG AAAATAA ATACAAA GGTCTT GAGGGT GTGTAA ATTGAAA GCGAGAA ATAATCA TAAATTA
TTTTAC TTTTAT TATGTTT CCAAGAA CTCCCA CACRATT TAACITT CGTCTT TATTAGT ATTATAT

30
H6 promoter
MetGlu GluProGln SerAsp ProSer ValGluPro
3151 TTTCATT ATCGCGA TATCCGT TAAGTTT GTATCGT AATGGAG GAGCCCG AGTCAGA TCCTAGC GTCGAGC
AAAGTAA TAGCGCT ATAGGCA ATTCAAA CATAGCA TTACCTC CTGGGG TCAGTCT AGGATCG CAGCTCG

35
p53
3221 ..ProLeu SerGln GluThrPhe SerAsp LeuTrp LysLeuLeu ProGlu AsnAsn ValLeuSer ProLeu
CCCCCT GAGTCAG GAAACAT TTTCAGA CCTATGG AACTAC TTCTGA AAACAAC GTTCTGT CCCCTT
GGGGAGA CTCAGTC CTTTGTA AAAGTCT GGATACC TTGTATG AAGGACT TTTGTG CAGACA GGGGGAA

40
p53
3291 .ProSer GlnAlaMet AspAsp LeuMet LeuSerPro AspAsp IleGlu GlnTrpPhe ThrGlu AspPro
GCCGTCC CAAGCAA TGGATGA TTGTATG CTGTCCC CGGACGA TATTGAA CAATGGT TCACGTA AGACCCA
CGGCAGG GTTCGTT ACCTACT AACTAC GACAGGG GCCTGCT ATAACIT GTTACCA AGTGACT TCTGGST

45
p53
3361 GlyProAsp GluAla ProArg MetProGlu AlaAla ProPro ValAlaPro AlaPro AlaAla ProThrPro
GGTCAG ATGAAGC TCCAGA ATGCCAG AGGCTGC TCCCCC GTGGCC CTGCACC AGCAGCT CCTACAC
CCAGGTC TACTTGG AGGGTCT TACGGTC TCCAGC AGGGGG CACGGG GACGTGG TCGTGA GGATGTG

50
p53
3431 ..AlaAla ProAla ProAlaPro SerTrp ProLeu SerSerSer ValPro SerGln LysThrTyr GlnGly
CGGGCG CCTGCA CCAGCCC CCTCTG GCCCTG TCATCTT CTGTCCC TCCAG AAAACCT ACCAGG
GCCGCC GGGACGT GGTGGG GGAGGAC CCGGGAC AGTAGAA GACAGG AAGGGT TTTTGA TGGTCCC

55
p53
3501 .SerTyr GlyPheArg LeuGly PheLeu HisSerGly ThrAla LysSer ValThrCys ThrTyr SerPro
CAGCTAC GGTTCCT GTCTGG CTCTTG CATCTG GGACAGC CAAGTCT GTGACTT GCACGTA CTCCTCT
GTCGATG CCAAAGG CAGACCC GAAGAAC GTAAGAC CCTGTGG GTTCAGA CACTGAA CGTGCA GAGGGGA

60
p53
3571 AlaLeuAsn LysMet PheCys GlnLeuAla LysThr CysPro ValGlnLeu TrpVal AspSer ThrProPro
GCCCTCA ACAAGAT GTTTGC CAACTGG CCAAGAC CTGCCCT GTGCAGC TGTGGGT TGATTCC ACACCCC
CGGGAGT TGTCTA CAAACG GTTGACC GGTCTG GACGGGA CACGTC ACACCCA ACTAAGG TGTGGG

65
p53
3641 ..ProGly ThrArg ValArgAla MetAla IleTyr LysGlnSer GlnHis MetThr GluValVal ArgArg
CGCCCG CACCGC GTCCCG CCATGG CATCTAC AAGCAGT CACAGCA CATGAC GAGGTG TGAGGCG
GCGGCC GTGGCG CAGGCG GTTACG GTAGATG TCGTCA GTGTCT GTACTGC CTCCAAC ACTCCG

70
p53
3711 .CysPro HisHisGlu ArgCys SerAsp SerAspGly LeuAla ProPro GlnHisLeu IleArg ValGlu
CTGCCCC CACCATG AGCGCTG CTCAGAT AGCGATG GTCTGGC CCTCTCT CAGCATC TTATCC AGTGGAA
GACGGGG GTGGTAC TCGCGAC GAGTCTA TCGCTAC CAGACCG GGGAGGA GTCGTAG AATAGGC TCACCTT

		p53
5	3781	GlyAsnLeu ArgVal GluTyr LeuAspAsp ArgAsn ThrPhe ArgHisSer ValVal ValPro TyrGluPro· GGAAATT TGCCTGT GGAGTAT TTGGATG ACAGAAA CACTTTT CGACATA GTGTGGT GGTGCC TATGAGC CCTTTAA ACGCACA CCTCATA AACCTAC TGTCTTT GTGAAAA GCTGTAT CACACCA CCACGGG ATACTCG
		p53
10	3851	..ProGlu ValGly SerAspCys ThrThr IleHis TyrAsnTyr MetCys AsnSer SerCysMet GlyGly· CGCCTGA GGTGGGC TCTGACT GTACCAC CATCCAC TACAAC TACATGT TAACAGT TCCTGCA TGGGCGG GCGGACT CCAACCG AGACTGA CATGGTG GTAGGTG ATGTTGA TGTACAC ATTGTCA AGGACGT ACCCGCC
		p53
15	3921	.MetAsn ArgArgPro IleLeu ThrIle IleThrLeu GluAsp SerSer GlyAsnLeu LeuGly ArgAsn CATGAAC CGGAGGC CCATCCT CACCATC ATCACAC TGGAGA CTCCAGT GGTAAAT TACTGGG ACGGAAC GTACTTG GCCTCCG GGTAGGA GTGGTAG TGTGTG ACCTTCT GAGGTCA CCATTAG ATGACCC TGCCTTG
		p53
20	3991	SerPheGlu ValArg ValCys AlaCysPro GlyArg AspArg ArgThrGlu GluGlu AsnLeu ArgLysLys· AGCTTTG AGGTGCG TGTTTGT GCCTGTC CTGGGAG AGACCGG CGCACAG AGGAAGA GAATCTC CGCAAGA TCGAAAC TCCACGC ACAACA CGGACAG GACCCCT TCTGGCC GCGTGC TCCTTCT CTAGAG GCGTTCT
		p53
25	4061	..GlyGlu ProHis HisGluLeu ProPro GlySer ThrLysArg AlaLeu ProAsn AsnThrSer SerSer· AAGGGGA GCCTCAC CACGAGC TGCCCCC AGGGAGC ACTAAGC GAGCACT GCCCAAC AACACCA GCTCCTC TTCCCTT CGGAGTG GTGCTCG ACGGGGG TCCCTCG TGATTG CTCGTGA CGGGTTG TTGTGGT CGAGGAG
		p53
30	4131	.ProGln ProLysLys LysPro LeuAsp GlyGluTyr PheThr LeuGln IleArgGly ArgGlu ArgPhe TCCCCAG CCAAGA AGAAACC ACTGGAT GGAGAAT ATTTTAC CCTTCAG ATCCGTG GCGGTGA GCGCTTC AGGGGTC GGTTTCT TCTTTG TGACCTA CCTCTTA TAAAGTG GGAAGTC TAGGCAC CGGCACT CGCGAAG
		p53
35	4201	GluMetPhe ArgGlu LeuAsn GluAlaLeu GluLeu LysAsp AlaGlnAla GlyLys GluPro GlyGlySer· GAGATGT TCCGAGA GCTGAAT GAGGCCT TGGAACT CAAGGAT GCCCAGG CTGGGAA GGAGCCA GGGGGGA CTCTACA AGGCTCT CACTTA CTCCGA ACCTTGA GTTCCTA CGGGTCC GACCCCT CCTCGGT CCCCCCT
		p53
40	4271	..ArgAla HisSer SerHisLeu LysSer LysLys GlyGlnSer ThrSer ArgHis LysLysLeu MetPhe· GCAGGGC TCACTCC AGCCACC TGAAGTC CAAAAG GGTCACT CTACCTC CGGCCAT AAAAACC TCATGTT CGTCCG AGTGAGG TCGGTGG ACTTCAG GTTTTTC CCACTCA GATGGAG GCGGTA TTTTGTG AGTACAA
		p53
45	4341	.LysThr GluGlyPro AspSer Asp*** CAAGACA GAAGGGC CTGACTC AGACTGA ACGCGTT TTTTATC CCGGGCT CGAGGGT ACOGGAT CCTTTT GTTCTGT CTTCCTG GACTGAG TCTGACT TGCGCAA AAAATAG GGCCTGA GCTCCCA TGGCCTA GGAAGAA
		p53
50	4411	ATAGCTA ATTAGTC ACGTACC TTTGAGA GTACACC TTCAGCT ACCTCTT TTGTGTC TCAGAGT AACTTTC TATCGAT TAATCAG TGCATGG AAATCTT CATGGTG AAGTCGA TGGAGAA AACACAG AGTCTCA TTGAAG
		p53
55	4481	TTTAATC AATTCCA AAACAGT ATATGAT TTTCCAT TTCTTTC AAAGATG TAGTTTA CATCTGC TCCTTTG AAATTAG TTAAGGT TTTGTCA TATACIA AAAGGTA AAGAAAG TTTCTAC ATCAAAT GTAGAGC AGGAAAC
		Right Arm
60	4551	TTGAAAA GTAGCCT GAGCACT TCTTTTC TACCATG AATTACA GCTGGCA AGATCAA TTTTCC CAGTTCT AACTTTT CATCGGA CTCGTGA AGAAAG ATGGTAC TTAATGT CGACCGT TCTAGTT AAAAAGG GTCAGAA
		Right Arm
65	4621	GGACATT TTATTTT TTTTAAG TAGTGTG CTACATA TTTCAAT ATTCCA GATTGTA CAGCGAT CATTAAA CCTGTAA AATAAAA AAAATTC ATCACAC GATGTAT AAAGTTA TAAAGGT CTAACAT GTCGCTA GTAATTT
		Right Arm
70	4691	GGAGTAC GTCCCAT GTTATCC AGCAAGT CAGTATC AGCACCT TTGTTCA ATAGAAG TTTAACC ATTGTTA CCTCATG CAGGGTA CAATAGG TCGTTCA GTCATAG TCGTGA AACAAAG TATCTTC AAATTGG TAACAA
		Right Arm
	4761	AATTTTT ATTTGAT ACGGCTA TATGTAG AGGAGTT AACCGAT CCGTGTG TGAATA TCTACAT CCGCGGA TTAAAAA TAAACTA TGCCGAT ATACATC TCCTCAA TTGGCTA GGCACAA ACTTTAT AGATGTA GCGGGCT
		Right Arm
	4831	ATGAGCC AATAGAA GTTTAAC CAATTA ACTTTGT TAAGGTA AGCTGCC AAACACA AAGGAGT AAAGCCT TACTCGG TTATCTT CAATTG GTTTAAT TGAACA ATTCCAT TCGACGG TTTGTGT TTCCTCA TTTGCGA
		Right Arm
	4901	CCGCTGT AAAGAAC ATTGITT ACATAGT TATTCTT CAACAGA TCITTCA CIATTTT GTAGTCG TCTCTCA GGCGACA TTTCTTG TAACAAA TGTATCA ATAAGAA GTTGTCT AGAAAGT GATAAAA CATCAGC AGAGAGT
		Right Arm
	4971	ACACCGC ATCATGC AGACAAG AAGTTGT GCATTCA GTAACCT CAGGTTT AGCTCCA TACCTCA TCAAGAT TGTGGCG TAGTACG TCTGTTC TTCAACA CGTAAGT CATTGAT GTCCAAA TCGAGGT ATGGAGT AGTTCTA
		Right Arm
	5041	TTTTATA GCCTCGG TATTCTT GAACATT ACAGCCA TTTCAG AGGAGAT TGTAGAG TACCATA TTCCGTG

		AAAAATAT	CGGAGCC	ATAAGAA	CITGTAA	TGTCGGT	AAAGTTC	TCCTCTA	ACATCTC	ATGGTAT	AAGGCAC
					Right Arm						
5	5111	TTAGGGT	CGAATCC	ATGTGCC	AAAAACC	TATTTAG	AGATGCA	TGTTCAT	TATCCAT	GATAGCC	TCACAGA
		AATCCCA	GCTTAGG	TAACAGG	TTTTTGG	ATAAATC	TCTACGT	AACAGTA	ATAGGTA	CTATCGG	AGTGTCT
					Right Arm						
	5181	CGTATAT	GTAAGCC	ATCTTGA	ATGTATA	ATTTTGT	TGTTTTT	AACAAAC	GCTCGTG	AACAGCT	TCTATAC
		GCATATA	CATTCGG	TAGAACT	TACATAT	TAAACA	ACAAAG	TGTTTGG	CGAGCAC	TTGTCGA	AGATATG
					Right Arm						
10	5251	TTTTTCA	TTTTCTT	CATGATT	AATATAG	TTTACGG	AATATAA	GTATACA	AAAAGTT	TATAGTA	ATCTCAT
		AAAAAGT	AAAAGAA	GTACTAA	TTATATC	AAATGCC	TTATATT	CATATGT	TTTTCAA	ATATCAT	TAGAGTA
					Right Arm						
	5321	AATATCT	GAACAC	ATACATA	AAACATG	GAAGAAAT	TACACGA	TGTCGTT	GAGATAA	ATGGCTT	TTTATTTG
		TTATAGA	CITTTGTG	TATGTAT	TTTGTAC	CTTCTTA	ATGTGCT	ACAGCAA	CTCTATT	TACCGAA	AAATAAC
					Right Arm						
15	5391	TCATAGT	TTACAAA	TTGCGAG	TAATCTT	CATCTTT	TACGAAT	ATTGCAG	AATCTGT	TTTATCC	AACCACT
		AGTATCA	AATGTTT	AAGCGTC	ATTAGAA	GTAGAAA	ATGCTTA	TAACGTC	TTAGACA	AAATAGG	TTGGTCA
					Right Arm						
	5461	GATTTTT	GTATAAT	ATAACTG	GTATCCT	ATCTTCC	GATAGAA	TGCTGTT	ATTTAAC	ATTTTTG	CACCTAT
		CTAAAAA	CATATTA	TATTGAC	CATAGGA	TAGAAGG	CTATCTT	ACGACAA	TAAATTG	TAAAAAC	GTGGATA
					Right Arm						
20	5531	TAAGTTA	CATCTGT	CAAATCC	ATCTTTC	CAACTGA	CTTTATG	TAACGAT	GCGAAAT	AGCATTT	ATCACTA
		ATTCAAT	GTAGACA	GTTTAGG	TAGAAAG	GTTGACT	GAAATAC	ATTGCTA	CGCTTTA	TCGTAAA	TAGTGAT
					Right Arm						
25	5601	TGTCGTA	CCCAATT	ATCATGA	CAAGATT	CTCTTAA	ATAOGTA	ATCTTAT	TATCTCT	TGCATAT	TCGTAAT
		ACAGCAT	GGGTAA	TAGTACT	GTTCTAA	GAGAATT	TATGCAT	TAGAATA	ATAGAGA	ACGTATA	AGCATTAA
					Right Arm						
	5671	AGTAATT	GTAAAGA	GTATACG	ATAACAG	TATAGAT	ATACACG	TGATATA	AATATTT	AACCCCA	TTCTCTGA
		TCATTAA	CATTTCT	CATATGC	TATTGTC	ATATCTA	TATGTGC	ACTATAT	TTATAAA	TTGGGGT	AAGGACT
					Right Arm						
30	5741	GTAAAT	AATTACG	ATATTAC	ATTTCTT	TTTATTA	TTTTTAT	GTITTAG	TTATTTG	TTAGGTT	ATACAAA
		CATTTTA	TTAATGC	TATAATG	TAAAGGA	AAATAAT	AAAAATA	CAAAATC	AATAAAC	AATCCAA	TATGTTT
					Right Arm						
	5811	AATTATG	TTTATTT	GTGTATA	TTTAAAG	CGTCGTT	AAGAATA	AGCTTAG	TTAACAT	ATTATCG	CTTAGGT
		TTAATAC	AAATAAA	CACATAT	AAATTTT	GCAGCAA	TTCTTAT	TCGAATC	AATTGTA	TAATAGC	GAATCCA
					Right Arm						
35	5881	TTTGTAG	TATTTGA	ATCCTTT	CTTTAAA	TGGATTA	TTTTTCC	AATGCAT	ATTIATA	GCTTCAT	CCAAAGT
		AAACATC	ATAAATC	TAGGAAA	GAAATTT	ACCTAAT	AAAAAGG	TTACGTA	TAAATAT	CGAAGTA	GGTTTCA
					Right Arm						
40	5951	ATAACAT	TTAACAT	TCAGAAAT	TGCGGCC	GCAATTC	AATTCGT	AATCATG	GTCCATG	CTGTTC	CTGTGTG
		TATTGTA	AATTGTA	AGTCTTA	ACGCCGG	CGTTAAG	TTAAGCA	TTAGTAC	CAGTATC	GACAAAG	GACACAC
					Right Arm						
45	6021	AAATTGT	TATCCGC	TCACAAAT	TCCACAC	AACATAC	GAGCCGG	AAGCATA	AAGTGTA	AAGCCTG	GGGTGCC
		TTTAAAC	ATAGGCG	AGTGTGA	AGGTGTG	TTGTATG	CTCGGCC	TTGTAT	TTCCAT	TTCCGAC	CCCAACG
	6091	TAAATGAG	TGAGCTA	ACTCACA	TTAATTG	CGTTGCG	CTCAGTG	CCCGCTT	TCCAGTC	GGGAACG	CTGTCTG
		ATTACTC	ACTCGAT	TGAGTGT	AATTAAC	GCAACGC	GAGTGAC	GGGCGAA	AGGTCAG	CCCTTTG	GACAGCA
	6161	GCCAGCT	GCATTAA	TGAATCG	GCCAACG	CGCGGGG	AGAGGCG	GTTTGGG	TATTTGG	CGCTCTT	CCGCTTC
		CGGTGCA	CGTAATT	ACTTAGC	CGGTGCG	GCGCCCC	TCTCCGC	CAACCGC	ATAACCC	GCGAGAA	GCGGAAG
	6231	CTCCGTC	ACTGACT	CGCTGCG	CTCGGTC	GTCGCGC	TGCGGCG	AGCGGTA	TCAGCTC	ACTCAAA	GGCGGTA
		GAGCGAG	TGACIGA	GCGACGC	GAGCCAG	CAAGCCG	ACGCCGC	TCGCCAT	AGTCGAG	TGAGTTT	CCGCCAT
	6301	ATACGGT	TATCCAC	AGAATCA	GGGGATA	ACGCAGG	AAAGAAC	ATGTGAG	CAAAAGG	CCAGCAA	AAGGCCA
		TATGCCA	ATAGGTG	TCTTAGT	CCCCTAT	TGCGTCC	TTTCTTG	TACACTC	GTTTTCC	GGTCGTT	TTCCGGT
	6371	GGAAACG	TAAARAG	GCCCGGT	TGCTGGC	GTTTTC	CATAGGC	TCGCCCC	CCCTGAC	GAGCATC	ACAAAAA
		CCTTGGC	ATTTTTT	CGCGGCA	ACGACCG	CAAAAGG	GTATCCG	AGGCGGG	GGGACTG	CTCGTAG	TGTTTTT
	6441	TGACGCG	TCAAGTC	AGAGGTG	GCGAAAC	CCGACAG	GACTATA	AAGATAC	CAGGCGT	TTCCCCC	TGGAAGC
		AGCTGCG	AGTTTCA	TCTCCAC	CGCTTTG	GGCTGTC	CTGATAT	TTCTATG	GTCCGCA	AAGGGGG	ACCTTCG
	6511	TCCCTCG	TGCGCTC	TCTGTT	CCGACCC	TGCGGCT	TACCGGA	TACCTGT	CCGCCTT	TCTCCCT	TCGGGAA
		AGGGAGC	ACGCGAG	AGGACAA	GGCTGGG	ACGGCGA	ATGGCCT	ATGGACA	GGCGGAA	AGAGGGA	AGCCCTT
	6581	GCGTGGC	GCTTTCT	CATAGCT	CACGCTG	TAGGTAT	CTCAGTT	CGGTGTA	GGTCGTT	CGCTCCA	AGCTGGG
		CGCACCG	CGAARAG	GTATCGA	GTGCGAC	ATCCATA	GAGTCAA	GCCACAT	CCAGCAA	GCGAGGT	TCGACCC
	6651	CTGTGTG	CACGAAC	CCCCCGT	TCAGCCC	GACCGCT	GCGCCTT	ATCCGGT	AACATATC	GTCTTGA	GTCCAAC
		GACACAC	GTGCTTG	GGGGGCA	AGTCGGG	CTGGCGA	CGCGGAA	TAGGCCA	TTGATAG	CAGAACT	CAGGTTG
	6721	CCGGTAA	GACACGA	CTTATCG	CCAATGG	CAGCAGC	CACCTGT	AACAGGA	TTAGCAG	AGCGAGG	TATGTAG
		GGCCATT	CTGTGCT	GAATAGC	GGTGACC	GTCGTGC	GTGACCA	TTGTCTT	AATCGTC	TCGTCC	ATACATC
	6791	GCGGTGC	TACAGAG	TTCTTGA	AGTGGTG	GCTTAAC	TACGGCT	ACACTAG	AAGGACA	GTATTTG	GTATCTG
		CGCCACG	ATGCTCT	AAGAACT	TCACCAAC	CGGATTG	ATGCGGA	TGTGATC	TTCTGT	CATAAAC	CATAGAC
	6861	CGCTCTG	CTGAAGC	CAGTTAC	CTTCGGA	AAAGAGG	TTGGTAG	CTCTTGA	TCCGGCA	AACAAAC	CACCGCT
		GCGAGAC	GACTTCG	GTCAATG	GAAGCCT	TTTTCTC	AACCATC	GAGAACT	AGGCCGT	TTGTTTG	GTGGCGA
	6931	GGTAGCG	GTGGTTT	TTTTTGT	TGCAAGC	AGCAGAT	TACGCGC	AGAAAAA	AAGGATC	TCAAGAA	GATCCTT
		CCATCGC	CACCAAA	AAAACAA	ACGTTCA	TCGTCTA	ATGCGCG	TCTTTTT	TTCTTAG	AGTTCTT	CTAGGAA
	7001	TGATCTT	TTCTAGC	GGGTCTG	ACGCTCA	GTGGAAC	GAAAACT	CACGTTA	AGGGATT	TTGGTCA	TGAGATT
		ACTAGAA	AAGATGC	CCCAGAC	TGCGAGT	CACCTTG	CTTTTGA	GTGCAAT	TCCCTAA	AACCACT	ACTCTAA
	7071	ATCAAAA	AGGATCT	TCACCTA	GATCCTT	TAAATTT	AAAAATG	AAGTTTT	AAATCAA	TCTAAGG	TATATAT
		TAGTTTT	TCCTAGA	AGTGGAT	CTAGGAA	AATTTAA	TTTTTAC	TTCAAAA	TTTAGTT	AGATTTC	ATATATA

```

7141  GAGTAAA CTTGGTC TGACAGT TACCAAT GCTTAAT CAGTGAG GCAOCTA TCTCAGC GATCTGT CTATTTC
      CTCATTT GAACCAG ACTGTCA ATGGTTA CGAATTA GTCACCT CGTGGAT AGAGTCG CTAGACA GATAAAG
      -----
5    7211  GTTCATC CATAGTT GCCTGAC TCCCGT CGTGTAG ATAAC TAACG GAGGGC TTACCAT CTGGCCC
      CAAGTAG GTATCAA CGGACTG AGGGGCA GCACATC TATTGAT GCTATGC CCTCCCG AATGGTA GACCGGG
      Amp resistance gene
10   7281  CAGTGCT GCAATGA TACCGCG AGACCCA CGCTCAC CGGCTCC AGATTTA TCAGCAA TAAACCA GCCAGCC
      GTCACGA CGTTACT ATGGCGC TCTGGGT GCGAGTG GCCGAGG TCTAAAT AGTCGTT ATTTGGT CGGTCGG
      Amp resistance gene
15   7351  GGAAGGG CCGAGCG CAGAAGT GGTCCGT CAACITTT ATCCGCC TCCATCC AGTCTAT TAATTGT TGCCGGG
      CCTTCCC GGCTCGC GTCTTCA CCAGGAC GTTGAAA TAGGCGG AGGTAGG TCAGATA ATTAACA ACGGCC
      Amp resistance gene
      7421  AAGCTAG AGTAAGT AGTTCGC CAGTTAA TAGTTTG CGCAACG TTGTTCG CATTCGT ACAGGCA TCGTGGT
15   TTCGATC TCATTCA TCAAGCG GTCAATT ATCAAAC GCGTTGC AACCAAG GTAACGA TGTCCGT AGCACCA
      Amp resistance gene
      7491  GTCAGCG TCGTCGT TTGGTAT GGCTTCA TTCAGCT CCGGTTT CCAACGA TCAAGGC GAGTTAC ATGATCC
      CAGTGCG AGCAGCA AACCATA CCGAAGT AAGTCGA GGCCAAAG GGTTCGT AGTTCCG CTCATG TACTAGG
      Amp resistance gene
20   7561  CCCATGT TGTGCAA AAAAGCG GTTAGCT CCTTCGG TCCTCCG ATCGTTG TCAGAAG TAAGTTG GCCGCAG
      GGGTACA ACACGTT TTTTCGC CAATCGA GGAAGCC AGGAGGC TAGCAAC AGTCCTC ATTCAAC CCGCGTC
      Amp resistance gene
      7631  TGTATAT ACTCATG GTTATGG CAGCACT GCATAAT TCTCTTA CTGTCAT GCCATCC GTAAGAT GCTTTTC
25   ACAATAG TGAGTAC CAATACC GTCGTGA CGTATTA AGAGAAT GACAGTA CGGTAGG CATTCTA CGAAAAAG
      Amp resistance gene
      7701  TGTGACT GGTGAGT ACTCAAC CAAGTCA TTCTGAG AATAGTG TATGCGG CGACCGA GTTGCTC TTGCCCG
      ACACIGA CCACTCA TGAGTTG GTTCAGT AAGACTC TTATCAC ATACGCC GCTGGCT CAACGAG AACGGGC
      Amp resistance gene
30   7771  GCGTCAA TACGGGA TAATACC GCGCCAC ATAGCAG AACTTTA AAAGTGC TCATCAT TGGAAAA CGTTCTT
      CGCAGTT ATGCCCT ATTATGG CCGCGTG TATCGTC TTGAAAT TTTCAGG AGTAGTA ACCTTTT GCAAGAA
      Amp resistance gene
      7841  CGGGGCG AAAACTC TCAAGGA TCTTACC GCTGTTG AGATCCA GTTOGAT GTAACCC ACTCGTG CACCCAA
      GCCCCGC TTTTGAG AGTTCCCT AGAATGG CGACAAC TCTAGGT CAAGCTA CATTTGG TGAGCAC GTGGGTT
      Amp resistance gene
35   7911  CTGATCT TCAGCAT CTTTAC TTTCACC AGCGTTT CTGGGTG AGCAAAA ACAGGAA GGCAAAA TGCCGCA
      GACTAGA AGTCGTA GAAAATG AAAGTGG TCGCAA GACCCAC TCGTTTT TGTCTT CCGTTTT ACGCGCT
      Amp resistance gene
      7981  AAAAAAG GAATAAG GCGGACA CGGAAAT GTTGAAT ACTCATA CTCTTCC TTTTTC ATATTAT TGAAGCA
40   TTTTTC CTTATTC CGCTGT GCCTTTA CAACTTA TGAGTAT GAGAAG AAAAAAG TATAATA ACTTCGT
      -----
      Amp resistance gene
      8051  TTTATCA GGGTTAT TGTCTCA TGAGCGG ATACATA TTTGAAT GTATTTA GAAAAAT AAACAAA TAGGGGT
      AAATAGT CCAATA ACAGAGT ACTCGCC TATGTAT AAACCTA CATAAAT CTTTTA TTGTTT ATCCCCA
45   8121  TCCGCGC ACATTTC CCCGAAA AGTGCCA CCTGACG TCTAAGA AACCATT ATTATCA TGACATT AACCTAT
      AGGCGCG TGTAAAG GGGCTTT TCACGGT GGAAGTG AGATTCT TTGGTAA TAATAGT ACTGTAA TTGGATA
      AAAAAATA GCGGTAT CACGAG
8191  TTTTAT CCGCATA GTGCTC

```

FIGURE 2A

		1				50
	mCEA (6D)	ATGGAGTCTC	CCTCGGCCCC	TCCCCACAGA	TGGTGCATCC	CCTGGCAGAG
5	mCEA (6D, 1st&2nd)	ATGGAGTCTC	CCTCGGCCCC	TCCCCACAGA	TGGTGCATCC	CCTGGCAGAG
		51				100
	mCEA (6D)	GCTCCTGCTC	ACAGCCTCAC	TTCTAACCTT	CTGGAACCCG	CCCACCACTG
	mCEA (6D, 1st&2nd)	GCTCCTGCTC	ACAGCCTCAC	TTCTAACCTT	CTGGAACCCG	CCCACCACTG
10		101				150
	mCEA (6D)	CCAAGCTCAC	TATTGAATCC	ACGCCGTTCA	ATGTCGCAGA	GGGGAAGGAG
	mCEA (6D, 1st&2nd)	CCAAGCTCAC	TATTGAATCC	ACGCCGTTCA	ATGTCGCAGA	GGGGAAGGAG
15		151				200
	mCEA (6D)	GTGCTTCTAC	TTGTCCACAA	TCTGCCCCAG	CATCTTTTTG	GCTACAGCTG
	mCEA (6D, 1st&2nd)	GTGCTTCTAC	TTGTCCACAA	TCTGCCCCAG	CATCTTTTTG	GCTACAGCTG
20		201				250
	mCEA (6D)	GTACAAAGGT	GAAAGAGTGG	ATGGCAACCG	TCAAATTATA	GGATATGTAA
	mCEA (6D, 1st&2nd)	GTACAAAGGT	GAAAGAGTGG	ATGGCAACCG	TCAAATTATA	GGATATGTAA
25		251				300
	mCEA (6D)	TAGGAACTCA	ACAAGCTACC	CCAGGGCCCCG	CATACAGTGG	TCGAGAGATA
	mCEA (6D, 1st&2nd)	TAGGAACTCA	ACAAGCTACC	CCAGGGCCCCG	CATACAGTGG	TCGAGAGATA
		301				350
	mCEA (6D)	ATATACCCCA	ATGCATCCCT	GCTGATCCAG	AACATCATCC	AGAATGACAC
	mCEA (6D, 1st&2nd)	ATATACCCCA	ATGCATCCCT	GCTGATCCAG	AACATCATCC	AGAATGACAC
30		351				400
	mCEA (6D)	AGGATTCTAC	ACCCTACACG	TCATAAAGTC	AGATCTTGTG	AATGAAGAAG
	mCEA (6D, 1st&2nd)	AGGATTCTAC	ACCCTACACG	TCATAAAGTC	AGATCTTGTG	AATGAAGAAG
35		401				450
	mCEA (6D)	CAACTGGCCA	GTTCGGGGTA	TACCCGGAGC	TGCCCAAGCC	CTCCATGTCC
	mCEA (6D, 1st&2nd)	CAACTGGCCA	GTTCGGGGTA	TACCCGGAA <u>C</u>	TCCCTAAGCC	TTCTAT <u>TAGC</u>
40		451				500
	mCEA (6D)	AGCAACA <u>ACT</u>	CCAAACCCGT	GGAGGACAAG	GATGCTGTGG	CCTTCACCTG
	mCEA (6D, 1st&2nd)	<u>TCCAATAATA</u>	<u>GTAAGCCTGT</u>	<u>CGAAGACAAA</u>	<u>GATGCCGT<u>CG</u></u>	<u>CTTTTACATG</u>
45		501				550
	mCEA (6D)	TGAACCTGAG	ACTCAGGACG	CAACCTACCT	GTGGTGGGTA	AACAATCAGA
	mCEA (6D, 1st&2nd)	<u>CGAGCCCGAA</u>	<u>ACTCAAGACG</u>	<u>CAACATATCT</u>	<u>CTGGTGGGTG</u>	<u>AACAACCAGT</u>
		551				600
	mCEA (6D)	GCCTCCCGGT	CAGTCCCAGG	CTGCAGCTGT	CCAATGGCAA	CAGGACCCTC
	mCEA (6D, 1st&2nd)	<u>CCCTGCCTGT</u>	<u>GTCCCTTAGA</u>	<u>CTCCA<u>ACTCA</u></u>	<u>GCAACGGAAA</u>	<u>TAGA<u>ACTCTG</u></u>
50		601				650
	mCEA (6D)	ACTCTATTCA	ATGTCACAAG	AAATGACACA	GCAAGCTACA	AATGTGAAAC
	mCEA (6D, 1st&2nd)	<u>ACCCTGT<u>TTA</u></u>	<u>ACGTGACCAG</u>	<u>GAACGACACA</u>	<u>GCAAGCTACA</u>	<u>AATGC<u>GAAAC</u></u>

FIGURE 2B

		651				700
	mCEA (6D)	CCAGAACCCA	GTGAGTGCCA	GGCGCAGTGA	TTCAGTCATC	CTGAATGTCC
5	mCEA (6D, 1st&2nd)	CCAAAATCCA	GTCAGCGCCA	GGAGGTCTGA	TTCAGTGATT	CTCAACGTGC
		701				750
	mCEA (6D)	TCTATGGCCC	GGATGCCCC	ACCATTTCCC	CTCTAAACAC	ATCTTACAGA
	mCEA (6D, 1st&2nd)	TTTACGGACC	CGATGCTCCT	ACAATCAGCC	CTCTAAACAC	AAGCTATAGA
10		751				800
	mCEA (6D)	TCAGGGGAAA	ATCTGAACCT	CTCCTGCCAC	GCAGCCTCTA	ACCCACCTGC
	mCEA (6D, 1st&2nd)	TCAGGGGAAA	ATCTGAATCT	GAGCTGTCT	GCCGCTAGCA	ATCCTCCCGC
15		801				850
	mCEA (6D)	ACAGTACTCT	TGGTTTGTC	ATGGGACTTT	CCAGCAATCC	ACCCAAGAGC
	mCEA (6D, 1st&2nd)	CCAATACAGC	TGGTTTGTC	ATGGCACTTT	CCAACAGTCC	ACCCAGGAAC
20		851				900
	mCEA (6D)	TCTTTATCCC	CAACATCACT	GTGAATAATA	GTGGATCCTA	TACGTGCCAA
	mCEA (6D, 1st&2nd)	TGTTCAATCC	CAATATTACC	GTGAACAATA	GTGGATCCTA	CACGTGCCAA
25		901				950
	mCEA (6D)	GCCCATAACT	CAGACACTGG	CCTCAATAGG	ACCACAGTCA	CGACGATCAC
	mCEA (6D, 1st&2nd)	GCTCAACAATA	GCGACACCGG	ACTCAACCGC	ACAACCGTGA	CGACGATTAC
30		951				1000
	mCEA (6D)	AGTCTATGAG	CCACCCAAAC	CCTTCATCAC	CAGCAACAAC	TCCAACCCCG
	mCEA (6D, 1st&2nd)	CGTGTATGAG	CCACCAAAAC	CATTCTAAC	TAGTAACAAT	TCTAACCCAG
35		1001				1050
	mCEA (6D)	TGGAGGATGA	GGATGCTGTA	GCCTTAACCT	GTGAACCTGA	GATTTCAGAAC
	mCEA (6D, 1st&2nd)	TTGAGGATGA	GGACGCACTT	GCATTAACTT	GTGAGCCAGA	GATTCAAAAT
40		1051				1100
	mCEA (6D)	ACAACCTACC	TGTGGTGGGT	AAATAATCAG	AGCCTCCCGG	TCAGTCCCAG
	mCEA (6D, 1st&2nd)	ACCACCTATT	TATGGTGGGT	CAATAACCAA	AGTTTGCCGG	TTAGCCCACG
45		1101				1150
	mCEA (6D)	GCTGCAGCTG	TCCAATGACA	ACAGGACCCT	CACTCTACTC	AGTGTACAAA
	mCEA (6D, 1st&2nd)	CTTGCACTTG	TCTAATGATA	ACCGCACATT	GACACTCCTG	TCCGTTACTC
50		1151				1200
	mCEA (6D)	GGAATGATGT	AGGACCCTAT	GAGTGTGGAA	TCCAGAACGA	ATTAAGTGTT
	mCEA (6D, 1st&2nd)	GCAATGATGT	AGGACCTTAT	GAGTGTGGCA	TTCAGAAATGA	ATTATCCGTT
		1201				1250
	mCEA (6D)	GACCACAGCG	ACCCAGTCAT	CCTGAATGTC	CTCTATGGCC	CAGACGACCC
	mCEA (6D, 1st&2nd)	GATCACTCCG	ACCCTGTTAT	CCTTAATGTT	TTGTATGGCC	CAGACGACCC
		1251				1300
	mCEA (6D)	CACCATTTCC	CCCTCATACA	CCTATTACCG	TCCAGGGGTG	AACCTCAGCC
	mCEA (6D, 1st&2nd)	AACTATATCT	CCATCATACA	CCTACTACCG	TCCGGCGTG	AACTTGAGCC

FIGURE 2C

		1301			1350	
	mCEA (6D)	TCTCCTGCCA	TGCAGCCTCT	AACCCACCTG	CACAGTATTC	TTGGCTGATT
5	mCEA (6D, 1st&2nd)	<u>TTTCT</u> <u>TGCCA</u>	<u>TGCAGC</u> <u>ATCC</u>	<u>AACCC</u> <u>CCCTG</u>	<u>CACAGT</u> <u>ACTC</u>	<u>CTGGCT</u> <u>GATT</u>
		1351			1400	
	mCEA (6D)	GATGGGAACA	TCCAGCAACA	CACACAAGAG	CTCTTTATCT	CCAACATCAC
	mCEA (6D, 1st&2nd)	<u>GATGG</u> <u>AAACA</u>	<u>TTCAGC</u> <u>AGCA</u>	<u>TACT</u> <u>CAAGAG</u>	<u>TTATT</u> <u>TTATAA</u>	<u>GCAACAT</u> <u>TAAC</u>
10		1401			1450	
	mCEA (6D)	TGAGAAGAAC	AGCGGACTCT	ATACCTGCCA	GGCCAATAAC	TCAGCCAGTG
	mCEA (6D, 1st&2nd)	<u>TGAGA</u> <u>AGAAC</u>	<u>AGCGG</u> <u>ACTCT</u>	<u>ATACT</u> <u>TGCCA</u>	<u>GGCCA</u> <u>ATAAC</u>	<u>TCAGC</u> <u>CCAGTG</u>
15		1451			1500	
	mCEA (6D)	GCCACAGCAG	GACTACAGTC	AAGACAATCA	CAGTCTCTGC	GGAGCTGCCC
	mCEA (6D, 1st&2nd)	<u>GTCAC</u> <u>AGCAG</u>	<u>GACTAC</u> <u>AGTT</u>	<u>AAA</u> <u>ACAATAA</u>	<u>CTGT</u> <u>TTCGC</u>	<u>GGAGC</u> <u>TGCCC</u>
20		1501			1550	
	mCEA (6D)	AAGCCCTCCA	TCTCCAGCAA	CAACTCCAAA	CCCGTGGAGG	ACAAGGATGC
	mCEA (6D, 1st&2nd)	<u>AAGCC</u> <u>CTCCA</u>	<u>TCTCC</u> <u>AGCAA</u>	<u>CAACT</u> <u>CCAAA</u>	<u>CCCGT</u> <u>GAGG</u>	<u>ACAAG</u> <u>GATGC</u>
25		1551			1600	
	mCEA (6D)	TGTGGCCTTC	ACCTGTGAAC	CTGAGGCTCA	GAACACAACC	TACCTGTGGT
	mCEA (6D, 1st&2nd)	<u>TGTGG</u> <u>CCTTC</u>	<u>ACCTG</u> <u>TGAAC</u>	<u>CTGAG</u> <u>GCTCA</u>	<u>GAACA</u> <u>CAACC</u>	<u>TACCT</u> <u>GTGGT</u>
30		1601			1650	
	mCEA (6D)	GGGTAAATGG	TCAGAGCCTC	CCAGTCAGTC	CCAGGCTGCA	GCTGTCCAAT
	mCEA (6D, 1st&2nd)	<u>GGGTAA</u> <u>ATGG</u>	<u>TCAGA</u> <u>GCCTC</u>	<u>CCAGT</u> <u>CAGTC</u>	<u>CCAGG</u> <u>CTGCA</u>	<u>GCTGT</u> <u>CCAAT</u>
35		1651			1700	
	mCEA (6D)	GGCAACAGGA	CCCTCACTCT	ATTCAATGTC	ACAAGAAATG	ACGCAAGAGC
	mCEA (6D, 1st&2nd)	<u>GGCAAC</u> <u>AGGA</u>	<u>CCCTC</u> <u>ACTCT</u>	<u>ATTCA</u> <u>ATGTC</u>	<u>ACAAG</u> <u>AAATG</u>	<u>ACGCA</u> <u>AGAGC</u>
40		1701			1750	
	mCEA (6D)	CTATGTATGT	GGAATCCAGA	ACTCAGTGAG	TGCAAACCGC	AGTGACCCAG
	mCEA (6D, 1st&2nd)	<u>CTATG</u> <u>TATGT</u>	<u>GGAAT</u> <u>CCAGA</u>	<u>ACTCA</u> <u>GTGAG</u>	<u>TGCAA</u> <u>ACCGC</u>	<u>AGTGAC</u> <u>CCAG</u>
45		1751			1800	
	mCEA (6D)	TCACCCCTGGA	TGTCCTCTAT	GGGCCGGACA	CCCCCATCAT	TTCCCCCCCA
	mCEA (6D, 1st&2nd)	<u>TCACCC</u> <u>TGGA</u>	<u>TGTCCT</u> <u>CTAT</u>	<u>GGGCC</u> <u>GGACA</u>	<u>CCCCC</u> <u>ATCAT</u>	<u>TTCCCC</u> <u>CCCA</u>
50		1801			1850	
	mCEA (6D)	GACTCGTCTT	ACCTTTTCGGG	AGCGGACCTC	AACCTCTCCT	GCCACTCGGC
	mCEA (6D, 1st&2nd)	<u>GACTCG</u> <u>TCTT</u>	<u>ACCTT</u> <u>TCGGG</u>	<u>AGCGG</u> <u>ACCTC</u>	<u>AACCT</u> <u>CTCCT</u>	<u>GCCACT</u> <u>CGGC</u>
55		1851			1900	
	mCEA (6D)	CTCTAACCCA	TCCCCGCAGT	ATTCTTGGCG	TATCAATGGG	ATACCGCAGC
	mCEA (6D, 1st&2nd)	<u>CTCTAA</u> <u>CCCA</u>	<u>TCCCC</u> <u>GCAGT</u>	<u>ATTCT</u> <u>TGGCG</u>	<u>TATCA</u> <u>ATGGG</u>	<u>ATACCG</u> <u>CAGC</u>
60		1901			1950	
	mCEA (6D)	AACACACACA	AGTTCTCTTT	ATCGCCAAAA	TCACGCCAAA	TAATAACGGG
	mCEA (6D, 1st&2nd)	<u>AACACA</u> <u>CACA</u>	<u>AGTTCT</u> <u>CTTT</u>	<u>ATCGC</u> <u>CAAAA</u>	<u>TCACG</u> <u>CCAAA</u>	<u>TAATA</u> <u>ACGGG</u>

FIGURE 2D

		1951		2000
	mCEA (6D)	ACCTATGCCT	GTTTGTCTC	TAACTTGGCT ACTGGCCGCA ATAATTCCAT
5	mCEA (6D, 1st&2nd)	ACCTATGCCT	GTTTGTCTC	TAACTTGGCT ACTGGCCGCA ATAATTCCAT
		2001		2050
	mCEA (6D)	AGTCAAGAGC	ATCACAGTCT	CTGCATCTGG AACTTCTCCT GGTCTCTCAG
10	mCEA (6D, 1st&2nd)	AGTCAAGAGC	ATCACAGTCT	CTGCATCTGG AACTTCTCCT GGTCTCTCAG
		2051		2100
	mCEA (6D)	CTGGGGCCAC	TGTCGGCATC	ATGATTGGAG TGCTGGTTGG GGTGCTCTG
	mCEA (6D, 1st&2nd)	CTGGGGCCAC	TGTCGGCATC	ATGATTGGAG TGCTGGTTGG GGTGCTCTG
15		2101		
	mCEA (6D)	ATATAG		
	mCEA (6D, 1st&2nd)	ATATAG		

FIGURE 3**A. Amino Acid Sequence Comparison of “Wild-Type KSA” (1) and Modified KSA (2)**

5 1 MAPPQVLAFGLLLAAATATFAAAQEECVCENYKLAVNCFVNNNRQCQCTSVGAQNTVIC
 2 MAPPQVLAFGLLLAAATATFAAAQEECVCENYKLAVNCFVNNNRQCQCTSVGAQNTVIC

 1 SKLAAKCLVMKAEMNGSKLGRRAKPEGALQNNGLYDPDCDESGLFKAKQCNGTSTCWC
 2 SKLAAKCLVMKAEMNGSKLGRRAKPEGALQNNGLYDPDCDESGLFKAKQCNGTSTCWC

10 1 VNTAGVRRTDKDEITCSERVITYWIIIELEKHKAREKPYDSKSLRTALQKEITTRYQLD
 2 VNTAGVRRTDKDEITCSERVITYWIIIELEKHKAREKPYDSKSLRTALQKEITTRYQLD

 1 PKFITSILYENNVTITDLVQNSSQKTQNDVDIADVAYYFEKDVKGESLFHSSKMDLTVN
 2 PKFITSILYENNVTITDLVQNSSQKTQNDVDIADVAYYFEKDVKGESLFHSSKMDLTVN

15 1 GEQLDLDPGQTLIYYVDEKAPEFSMQGLKAGVIAVIVVVVIAVVAGIVVLVISRKKRMA
 2 GEQLDLDPGQTLIYYVDEKAPEFSMQGLKAGVIAVIVVVVIAVVAGIVVLVISRKKRMA

20 1 KYEKAIEIKEMGEMHRELNA
 2 KYEKAIEIKEMGEMHRELNA

B. DNA Sequence of Modified KSA

atggcgccccgcaggtcctcgcttcgggcttctgcttgccgcggcgacggcgacttttgccgcagctcaggaa

25 gaatgtgtctgtgaaaactacaagctggccgtaaactgctttgtgaataataatcgatcaatgccagtgacttca
 gttggtgcacaaaatactgtcatttgctcaaagctggctgccaaatgtttggtgatgaaggcagaaatgaatggc
 tcaaaacttgggagaagagcaaaacctgaaggggccctccagaacaatgatgggctttatgatcctgactgcat
 gagagcgggctctttaaggccaagcagtgcaacggcacctccacgtgctggtgtgtgaacactgctggggtcaga
 agaacagacaaggacactgaaataacctgctctgagcagtgagaacctactggatcatcattgaactaaacac

30 aaagcaagagaaaaaccttatgatagtaaaagtgttcgggactgcacttcagaaggagatcacaacgcgttatcaa
 ctggatccaaaatttatcacgagtgtgtgtatgagaataatgttatcactattgatctggttcaaaattcttct
 caaaaaactcagaatgatgtggacatagctgatgtggcttattattttgaaaagatgttaaaggatgaatccttg
 tttcatttctaagaaaatggacctgacagtaaatggggaacaactggatctggatcctggtcaaactttaatttat
 tatgttgatgaaaaagcacctgaattctcaatgcaggggtctaaaagctgggttattgctgttattgtggttg

35 gtgatagcagttgttgctggaattgttgctggttatttccagaaagaagagaatggcaaagtatgagaaggct
 gagataaaggagatgggtgagatgcatagggaaactcaatgcataa

FIGURE 4A
Construction of Modified KSA Plasmid

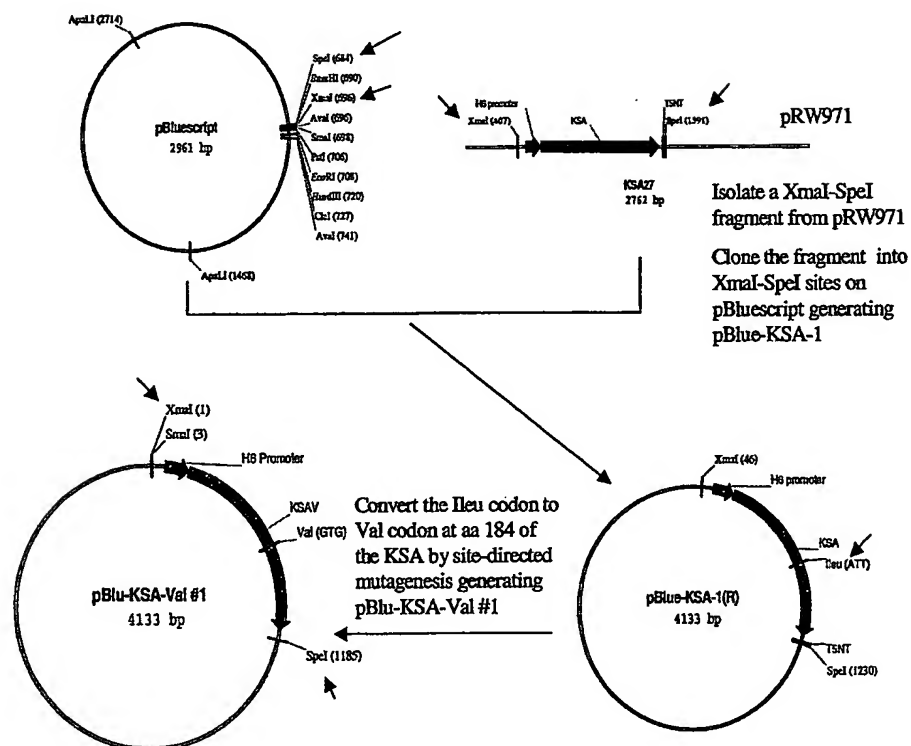


FIGURE 4B
Construction of Modified KSA Plasmid

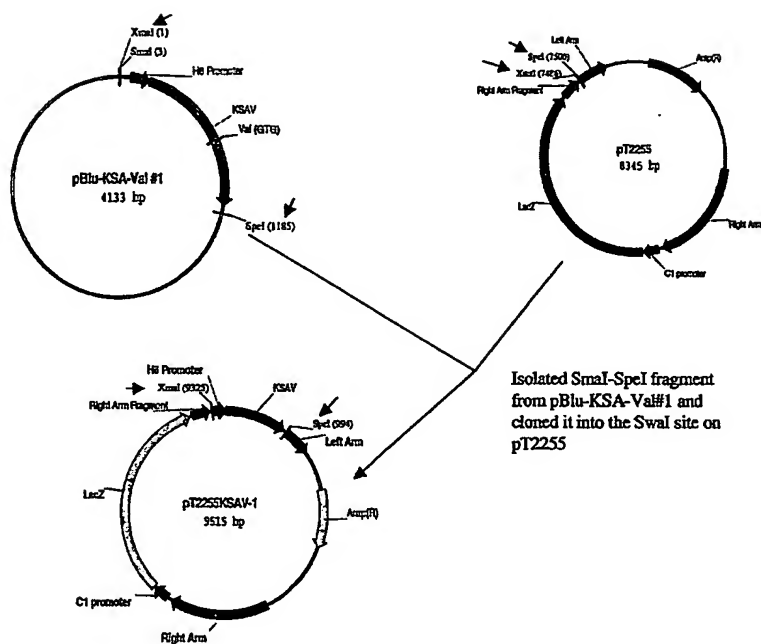
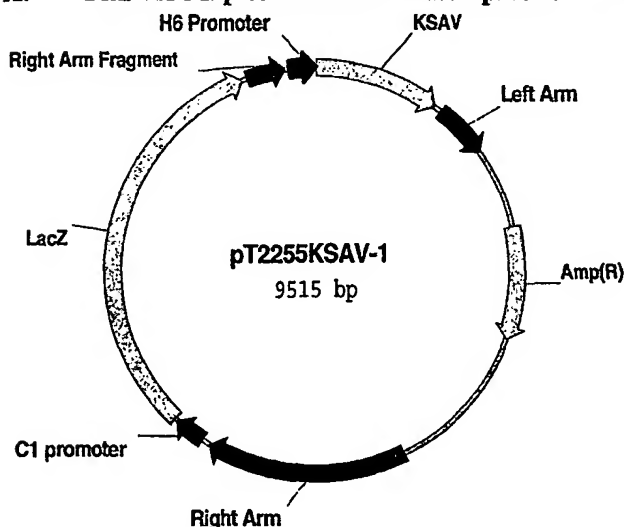


FIGURE 5**A. Plasmid Map of Modified KSA Expression Vector****B. DNA Sequence of Modified KSA Expression Vector**

Promoter H6 for KSAV	9930-9515
KSAV	1-945
Left arm	1002-1422
Right arm	4070-5590
Right arm fragment	9012-9299

1 MetAlaProPro GlnValLeu AlaPheGly LeuLeuLeuAla AlaAlaThr-
 5 ATGGCGCCCC CGCAGGTCCT CGCGTTCGGG CTTCTGCTTG CCGCGGCGAC
 10 TACCGCGGGG GCGTCCAGGA GCGCAAGCCC GAAGACGAAC GGCGCCGCTG
 15 .AlaThrPhe AlaAlaAlaGln GluGluCys ValCysGlu AsnTyrLysLeu-
 20 51 GGCGACTTTT GCCGAGCTC AGGAAGAATG TGTCTGTGAA AACTACAAGC
 CCGCTGAAAA CGGCGTCGAG TCCTTCTTAC ACAGACACTT TTGATGTTCTG
 25 ..AlaValAsn CysPheVal AsnAsnAsnArg GlnCysGln CysThrSer
 30 101 TGGCCGTAAA CTGCTTTGTG AATAATAATC GTCAATGCCA GTGTACTTCA
 ACCGGCATTG GACGAAACAC TTATTATTAG CAGTTACGGT CACATGAAGT
 35 ValGlyAlaGln AsnThrVal IleCysSer LysLeuAlaAla LysCysLeu-
 40 151 GTTGGTGAC AAAAATACTGT CATTGTGCTA AAGCTGGCTG CCAAATGTTT
 CAACCACGTG TTTTATGACA GTAAACGAGT TTCGACCGAC GGTTTACAAA
 45 .ValMetLys AlaGluMetAsn GlySerLys LeuGlyArg ArgAlaLysPro-
 50 201 GGTGATGAAG GCAGAAATGA ATGGCTCAAA ACTTGGGAGA AGAGCAAAAC
 CCACTACTTC CGTCTTTACT TACCGAGTTT TGAACCCTCT TCTCGTTTTG
 55 ..GluGlyAla LeuGlnAsn AsnAspGlyLeu TyrAspPro AspCysAsp
 60 251 CTGAAGGGGC CCTCCAGAAC AATGATGGGC TTTATGATCC TGACTGCGAT
 GACTTCCCCG GGAGGTCTTG TTAATAACCG AAATACTAGG ACTGACGCTA
 65 GluSerGlyLeu PheLysAla LysGlnCys AsnGlyThrSer ThrCysTrp-
 70 301 GAGAGCGGGC TCTTTAAGGC CAAGCAGTGC AACGGCACCT CCACGTGCTG
 CTCTCGCCCC AGAAATTCGG GTTCGTCACG TTGCCGTGGA GGTGCACGAC
 75 .CysValAsn ThrAlaGlyVal ArgArgThr AspLysAsp ThrGluIleThr-
 80 351 GTGTGTGAAC ACTGCTGGGG TCAGAAGAAC AGACAAGGAC ACTGAAATAA
 CACACACTTG TGACGACCCC AGTCTTCTTG TCTGTTCTTG TGACTTTATT
 85 ..CysSerGlu ArgValArg ThrTyrTrpIle IleIleGlu LeuLysHis

401 CCTGCTCTGA GCGAGTGAGA ACCTACTGGA TCATCATTGA ACTAAAACAC
 GGACGAGACT CGCTCACTCT TGGATGACCT AGTAGTAACT TGATTTTGTG
 LysAlaArgGlu LysProTyr AspSerLys SerLeuArgThr AlaLeuGln.
 451 AAAGCAAGAG AAAAACCTTA TGATAGTAAA AGTTTGCGGA CTGCACCTTCA
 5 TTTTCGTTCTC TTTTGGGAAT ACTATCATTT TCAAACGCCT GACGTGAAGT
 .LysGluIle ThrThrArgTyr GlnLeuAsp ProLysPhe IleThrSerVal.
 501 GAAGGAGATC ACAACGCGTT ATCAACTGGA TCCAAAATTT ATCACGAGTG
 CTTCTCTAG TGTTCGCAA TAGTTGACCT AGGTTTTTAA TAGTGCTCAC
 ..LeuTyrGlu AsnAsnVal IleThrIleAsp LeuValGln AsnSerSer
 10 551 TGTGTATGA GAATAATGTT ATCACTATTG ATCTGGTTCA AAATCTTCT
 ACAACATACT CTTATTACAA TAGTGATAAC TAGACCAAGT TTTAAGAAGA
 GlnLysThrGln AsnAspVal AspIleAla AspValAlaTyr TyrPheGlu.
 601 CAAAAAAGT AGAATGATGT GGACATAGCT GATGTGGCTT ATTATTTTGA
 GTTTTTTGTG TCTTACTACA CCTGTATCGA CTACACCGAA TAATAAACT
 15 .LysAspVal LysGlyGluSer LeuPheHis SerLysLys MetAspLeuThr.
 651 AAAAGATGTT AAAGGTGAAT CCTTGTTCCTA TTCTAAGAAA ATGGACCTGA
 TTTTCTACAA TTTCCACTTA GGAACAAAGT AAGATTCTTT TACCTGGACT
 ..ValAsnGly GluGlnLeu AspLeuAspPro GlyGlnThr LeuIleTyr
 701 CAGTAAATGG GGAACAACTG GATCTGGATC CTGGTCAAAC TTTAATTTAT
 20 GTCATTACC CCTTGTGAC CTAGACCTAG GACCAAGTTG AAATTAATA
 TyrValAspGlu LysAlaPro GluPheSer MetGlnGlyLeu LysAlaGly.
 751 TATGTTGATG AAAAAGCACC TGAATTCTCA ATGCAGGGTC TAAAAGCTGG
 ATACAACTAC TTTTTCGTGG ACTTAAGAGT TACGTCCCAG ATTTTCGACC
 .ValIleAla ValIleValVal ValValIle AlaValVal AlaGlyIleVal.
 801 TGTATTGCT GTTATTGTGG TTGTGGTGAT AGCAGTTGTT GCTGGAATTG
 25 ACAATAACGA CAATAACACC AACACCATA TCGTCAACAA CGACCTAAC
 ..ValLeuVal IleSerArg LysLysArgMet AlaLysTyr GluLysAla
 851 TTGTGCTGGT TATTTCCAGA AAGAAGAGAA TGGCAAAGTA TGAGAAGGCT
 AACACGACCA ATAAAGGTCT TTCTTCTCTT ACCGTTTCAT ACTCTCCGA
 30 GluIleLysGlu MetGlyGlu MetHisArg GluLeuAsnAla ***
 901 GAGATAAAGG AGATGGGTGA GATGCATAGG GAACTCAATG CATAAGAAGC
 CTCTATTTCC TCTACCACT CTACGTATCC CTTGAGTTAC GTATTCTTCG
 951 TTATCGATAC CGTCGACCTC GAGGAATTCT TTTTATTGAT TAACTAGTTA
 AATAGCTATG GCAGCTGGAG CTCCTTAAGA AAAATAACTA ATTGATCAAT
 35 1001 ATCACGCCCG CTTATAAAGA TCTAAAATGC ATAATTCTA AATAATGAAA
 TAGTGCCGGC GAATATTTCT AGATTTTACG TATTAAAGAT TTATTACTTT
 1051 AAAAAGTACA TCATGAGCAA CGCGTTAGTA TATTTTACAA TGGAGATTAA
 TTTTTCATGT AGTACTCGTT GCGCAATCAT ATAAATGTT ACCTCTAATT
 1101 CGCTCTATAC CGTTCTATGT TTATTGATTC AGATGATGTT TTAGAAAAGA
 40 GCGAGATATG GCAAGATACA AATAACTAAG TCTACTACAA AATCTTTTCT
 1151 AAGTTATTGA ATATGAAAAC TTTAATGAAG ATGAAGATGA CGACGATGAT
 TTCAATAACT TATACTTTTG AAATTACTTC TACTTCTACT GCTGCTACTA
 1201 TATTGTTGTA AATCTGTTT AGATGAAGAA GATGACGCGC TAAAGTATAC
 ATAACAACAT TTAGACAAA TCTACTTCTT CTACTGCGC ATTTTCATATG
 45 1251 TATGGTTACA AAGTATAAGT CTATACTACT AATGGCGACT TGTGCAAGAA
 ATACCAATGT TTCATATTCA GATATGATGA TTACCGCTGA ACACGTTCTT
 1301 GGTATAGTAT AGTGAAAATG TTGTTAGATT ATGATTATGA AAAACCAAAT
 CCATATCATA TCACTTTTAC AACAACTAA TACTAATACT TTTTGGTTTA
 1351 AAATCAGATC CATATCTAAA GGTATCTCCT TTGCACATAA TTTCATCTAT
 50 TTTAGTCTAG GTATAGATT CCATAGAGGA AACGTGTATT AAAGTAGATA
 1401 TCCTAGTTTA GAATACCTGC AGCCAAGCTT GGCACGCGC GTCGTTTTAC
 AGGATCAAAT CTTATGGACG TCGGTTGAA CCGTGACCGG CAGCAAAATG
 1451 AACGTCGTGA CTGGGAAAAC CCTGGCGTTA CCCAACTTAA TCGCCTTGCA
 TTGCAGCACT GACCCTTTTG GGACCGCAAT GGGTTGAATT AGCGGAACGT
 55 1501 GCACATCCCC CTTTCGCCAG CTGGCGTAAT AGCGAAGAGG CCCGCACCGA
 CGTGTAGGGG GAAAGCGGTC GACCGCATT TCGCTTCTCC GGGCGTGGCT
 1551 TCGCCCTTCC CAACAGTTGC GCAGCCTGAA TGGCGAATGG CGCCTGATGC
 AGCGGGAAGG GTTGTCAACG CGTCGCACTT ACCGCTTACC GCGGACTACG
 1601 GGTATTTTCT CCTTACGCAT CTGTGCGGTA TTTCACACCG CATATGGTGC

		CCATAAAAGA	GGAATGCGTA	GACACGCCAT	AAAGTGTCGC	GTATACCACG
1651		ACTCTCAGTA	CAATCTGCTC	TGATGCCGCA	TAGTTAAGCC	AGCCCCGACA
		TGAGAGTCAT	GTTAGACGAG	ACTACGGCGT	ATCAATTCCG	TCGGGGCTGT
1701		CCCGCCAACA	CCCGCTGACG	CGCCCTGACG	GGCTTGTCTG	CTCCCGGCAT
5		GGGCGGTTGT	GGGCGACTGC	GCGGGACTGC	CCGAACAGAC	GAGGGCCGTA
1751		CCGCTTACAG	ACAAGCTGTG	ACCGTCTCCG	GGAGCTGCAT	GTGTCAGAGG
		GGCGAATGTC	TGTTTCGACAC	TGGCAGAGGC	CCTCGACGTA	CACAGTCTCC
1801		TTTTTACCCT	CATCACCAGAA	ACGCGCGAGA	CGAAAGGGCC	TCGTGATACG
		AAAAGTGGCA	GTAGTGGCTT	TGCGCGCTCT	GCTTTCCCGG	AGCAGTATGC
10	1851	CCTATTTTTA	TAGGTTAATG	TCATGATAAT	AATGGTTTCT	TAGACGTCAG
		GGATAAAAAT	ATCCAATTAC	AGTACTATTA	TTACCAAAGA	ATCTGCAGTC
1901		GTGGCACTTT	TCGGGGAAAT	GTGCGCGGAA	CCCCTATTTG	TTTATTTTTT
		CACCGTGAAA	AGCCCCTTTA	CACGCGCCTT	GGGGATAAAC	AAATAAAAAG
1951		TAAATACATT	CAAAATATGT	TCCGCTCATG	AGACAATAAC	CCTGTATAAT
15		ATTTATGTAA	GTTTATACAT	AGGCGAGTAC	TCTGTTATTG	GGACTATTTA
2001		GCTTCAATAA	TATTGAAAAA	GGAAGAGTAT	GAGTATTCAA	CATTTCCGTG
		CGAAGTTATT	ATAACTTTTT	CCTTCTCATA	CTCATAAGTT	GTAAAGGCAC
2051		TCGCCCTTAT	TCCCTTTTTT	GCGGCATTTT	GCCTTCCTGT	TTTTGCTCAC
		AGCGGGAATA	AGGGAAAAAA	CGCCGTAAAA	CGGAAGGACA	AAAACGAGTG
20	2101	CCAGAAACGC	TGGTGAAAGT	AAAAGATGCT	GAAGATCAGT	TGGGTGCACG
		GGTCTTTGCG	ACCACTTTCA	TTTTCTACGA	CTTCTAGTCA	ACCCACGTGC
2151		AGTGGGTTAC	ATCGAACTGG	ATCTCAACAG	CGGTAAAGATC	CTTGAGAGTT
		TCACCCAATG	TAGCTTGACC	TAGAGTTGTC	GCCATTCTAG	GAAGTCTCAA
2201		TTCGCCCCGA	AGAACGTTTT	CCAATGATGA	GCATTTTTAA	AGTTCTGCTA
25		AAGCGGGGCT	TCTTGCAAAA	GGTTACTACT	CGTGAAAAAT	TCAAGACGAT
2251		TGTGGCGCGG	TATTATCCCG	TATTGACGCC	GGGCAAGAGC	AACTCGGTGC
		ACACCGCGCC	ATAATAGGGC	ATAACTGCGG	CCCGTTCTCG	TTGAGCCAGC
2301		CCGCATACAC	TATTCTCAGA	ATGACTTGGT	TGAGTACTCA	CCAGTCACAG
		GGCGTATGTG	ATAAGAGTCT	TACTGAACCA	ACTCATGAGT	GGTCAGTGTC
30	2351	AAAAGCATCT	TACGGATGGC	ATGACAGTAA	GAGAATTATG	CAGTGCTGCC
		TTTTTCGTAGA	ATGCCTACCG	TACTGTCAAT	CTCTTAATAC	GTACAGCGG
2401		ATAACCATGA	GTGATAACAC	TGCGGCCAAC	TTACTTCTGA	CAACGATCGG
		TATTGGTACT	CACTATTGTG	ACGCCGGTTG	AATGAAGACT	GTTGCTAGCC
2451		AGGACCGAAG	GAGCTAACCG	CTTTTTTGCA	CAACATGGGG	GATCATGTAA
35		TCCTGGCTTC	CTCGATTGGC	GAAAAAACGT	GTTGTACCCC	CTAGTACATT
2501		CTCGCCTTGA	TCGTTGGGAA	CCGGAGCTGA	ATGAAGCCAT	ACCAAACGAC
		GATCGGAAC	AGCAACCCCT	GGCCTCGACT	TACTTCGGTA	TGGTTTGCTG
2551		GAGCGTGACA	CCACGATGCC	TGTAGCAATG	GCAACAACGT	TGCGCAAACT
		CTCGCACTGT	GGTGCTACGG	ACATCGTTAC	CGTTGTGCA	ACGCGTTTGA
40	2601	ATTAATGGC	GAAGTACTTA	CTCTAGCTTC	CCGGCAACAA	TTAATAGACT
		TAATTGACCG	CTTGATGAAT	GAGATCGAAG	GGCCGTGTGT	AATTATCTGA
2651		GGATGGAGGC	GGATAAAGTT	GCAGGACCAC	TTCTGCGCTC	GGCCCTTCCG
		CCTACCTCCG	CCTATTTCAA	CGTCTGGTG	AAGACGCGAG	CCGGGAAGGC
2701		GCTGGCTGGT	TTATGCTGTA	TAAATCTGGA	GCCGCTGAGC	GTGGGTCTCG
45		CGACCGACCA	AATAACGACT	ATTTAGACCT	CGGCCACTCG	CACCCAGAGC
2751		CGGTATCATT	GCAGCACTGG	GGCCAGATGG	TAAGCCCTCC	CGTATCGTAG
		GCCATAGTAA	CGTCGTGACC	CCGGTCTACC	ATTCGGGAGG	GCATAGCATC
2801		TTATCTACAC	GACGGGGAGT	CAGGCAACTA	TGGATGAACG	AAATAGACAG
		AATAGATGTG	CTGCCCCCTA	GTCCGTTGAT	ACCTACTTGC	TTTATCTGTC
50	2851	ATCGCTGAGA	TAGGTGCCTC	ACTGATTAAG	CATTGGTAAC	TGTCAGACCA
		TAGCGACTCT	ATCCACGGAG	TGACTAATTC	GTAACCATTC	ACAGTCTGGT
2901		AGTTTACTCA	TATATACTTT	AGATTGATTT	AAAACCTCAT	TTTTAATTTA
		TCAAATGAGT	ATATATGAAA	TCTAACTAAA	TTTTGAAGTA	AAAATTAAT
2951		AAAGGATCTA	GGTGAAGATC	CTTTTTGATA	ATCTCATGAC	CAAAATCCCT
55		TTTCCTAGAT	CCACTTCTAG	GAAAAACTAT	TAGAGTACTG	GTTTTAGGGA
3001		TAACGTGAGT	TTTCGTTCCA	CTGAGCGTCA	GACCCCGTAG	AAAAGATCAA
		ATTGCACTCA	AAAGCAAGGT	GACTCGCAGT	CTGGGGCATC	TTTCTAGTT
3051		AGGATCTTCT	TGAGATCCTT	TTTTTCTGCG	CGTAATCTGC	TGCTTGCAAA
		TCCTAGAAGA	ACTCTAGGAA	AAAAAGACGC	GCATTAGACG	ACGAACGTTT

	3101	CAAAAAAACC	ACCGCTACCA	GCGGTGGTTT	GTTTGCCGGA	TCAAGAGCTA
		GTTTTTTTGG	TGGCGATGGT	CGCCACCAAA	CAAACGGCCT	AGTTCTCGAT
	3151	CCAACTCTTT	TTCCGAAGGT	AACTGGCTTC	AGCAGAGCGC	AGATACCAAA
		GGTTGAGAAA	AAGGCTTCCA	TTGACCGAAG	TCGTCTCGCG	TCTATGGTTT
5	3201	TACTGTCTTT	CTAGTGTAGC	CGTAGTTAGG	CCACCACTTC	AAGAACTCTG
		ATGACAGGAA	GATCACATCG	GCATCAATCC	GGTGGTGAAG	TTCTTGAGAC
	3251	TAGCACCGCC	TACATACCTC	GCTCTGCTAA	TCCGTGTACC	AGTGGCTGCT
		ATCGTGGCGG	ATGTATGGAG	CGAGACGATT	AGGACAATGG	TCACCGACGA
	3301	GCCAGTGGCG	ATAAGTCGTG	TCTTACCGGG	TTGGACTCAA	GACGATAGTT
10		CGGTACCCGC	TATTCAGCAC	AGAATGGCCC	AACCTGAGTT	CTGCTATCAA
	3351	ACCGGATAAG	GCGCAGCGGT	CGGGCTGAAC	GGGGGGTTCC	TGCACACAGC
		TGGCCTATTG	CGCGTCGCCA	GCCCGACTTG	CCCCCAAGC	ACGTGTGTCC
	3401	CCAGCTTTGA	GCGAACGACC	TACACCGAAC	TGAGATACCT	ACAGCGTGAG
		GGTCGAACCT	CGCTTGCTGG	ATGTGGCTTG	ACTCTATGGA	TGTCGCACTC
15	3451	CTATGAGAAA	GCGCCACGCT	TCCCGAAGGG	AGAAAGGCGG	ACAGGTATCC
		GATACTCTTT	CGCGGTGCGA	AGGGCTTCCC	TCTTTCGCC	TGTCCATAGG
	3501	GGTAAGCGGC	AGGGTCGGAA	CAGGAGAGCG	CACGAGGGAG	CTTCCAGGGG
		CCATTCGCCG	TCCCAGCCTT	GTCTCTCGC	GTGCTCCCTC	GAAGGTCCCC
	3551	GAAACGCCTG	GTATCTTTAT	AGTCTGTGCG	GGTTTCGCCA	CCTCTGACTT
20		CTTTGCGGAC	CATAGAAATA	TCAGGACAGC	CCAAAGCGGT	GGAGACTGAA
	3601	GAGCGTCGAT	TTTTGTGATG	CTCGTCAGGG	GGGCGGAGCC	TATGGAAAAA
		CTCGCAGCTA	AAAACACTAC	GAGCAGTCCC	CCCGCCTCGG	ATACCTTTTT
	3651	CGCCAGCAAC	GCGGCCTTTT	TACGGTTCCT	GGCCTTTTGC	TGGCCTTTTG
		GCGGTGCTTG	CGCCGAAAAA	ATGCCAAGGA	CCGAAAAACG	ACCGGAAAAAC
25	3701	CTCATATGTT	CTTTCCTGCG	TTATCCCCTG	ATTCTGTGGA	TAACCGTATT
		GAGTGTAACA	GAAAGGACGC	AATAGGGGAC	TAAGACACCT	ATTGGCATAA
	3751	ACCGCCTTTG	AGTGAGCTGA	TACCGCTCGC	CGCAGCCGAA	CGACCGAGCG
		TGGCGGAAAC	TCACTCGACT	ATGGCGAGCG	GCGTCGGCTT	GCTGGCTCGC
	3801	CAGCGAGTCA	GTGAGCGAGG	AAGCGGAAGA	GCGCCCAATA	CGCAAACCGC
30		GTGCGTCAGT	CACTCGCTCC	TTGCGCTTCT	CGCGGGTTAT	GCGTTTGGCG
	3851	CTCTCCCCGC	GCGTTGGCCG	ATTCAATTAAT	GCAGCTGGCA	CGACAGGTTT
		GAGAGGGGCG	CGCAACCGGC	TAAGTAATTA	CGTCGACCGT	GCTGTCCAAA
	3901	CCCGACTGGA	AAGCGGGCAG	TGAGCGCAAC	GCAATTAATG	TGAGTTAGCT
		GGGTGACCT	TTGCCCCGTC	ACTCGCGTTG	CGTTAATTAC	ACTCAATCGA
35	3951	CACTCATTAG	GCACCCCAGG	CTTTACACTT	TATGCTTCCG	GCTCGTATGT
		GTGAGTAATC	CGTGGGGTCC	GAAATGTGAA	ATACGAAGGC	CGAGCATACA
	4001	TGTTGTGGAAT	TGTGAGCGGA	TAACAATTTT	ACACAGGAAA	CAGCTATGAC
		ACACACCTTA	ACACTCGCCT	ATTGTTAAAG	TGTGTCCTTT	GTGATACTG
	4051	CATGATTACG	AATTGAATTG	CGGCCGCAAT	TCTGAATGTT	AAATGTTATA
40		GTACTAATGC	TTAACTTAAC	GCCGGCGTTA	AGACTTACAA	TTTACAATAT
	4101	CTTTGGATGA	AGCTATAAAT	ATGCATTGGA	AAAAATAATC	ATTTAAAGAA
		GAAACCTACT	TCGATATTTA	TACGTAACCT	TTTTATTAGG	TAAATTTCTT
	4151	AGGATTCAAA	TACTACAAAA	CCTAAGCGAT	AATATGTTAA	CTAAGCTTAT
		TCCTAAGTTT	ATGATGTTTT	GGATTGCGTA	TTATACAATT	GATTGCAATA
45	4201	TCTTAACGAC	GCTTTAAATA	TACACAAATA	AACATAATTT	TTGTATAACC
		AGAAATGCTG	CGAAATTTAT	ATGTGTTTAT	TTGTATTAAA	AACATATTGG
	4251	TAACAAATAA	CTAAAAACATA	AAAATAATAA	AAGGAAATGT	AATATCGTAA
		ATTGTTTATT	GATTTTGTAT	TTTTATTATT	TTCTTTTACA	TTATAGCATT
	4301	TTATTTTACT	CAGGAATGGG	GTTAAATATT	TATATCACGT	GTATATCTAT
50		AATAAAATGA	GTCTTATCCC	CAATTTATAA	ATATAGTGCA	CATATAGATA
	4351	ACTGTTATCG	TATACTCTTT	ACAATTACTA	TTACGAATAT	GCAAGAGATA
		TGACAAATAGC	ATATGAGAAA	TGTTAATGAT	AATGCTTATA	CGTTCTCTAT
	4401	ATAAGATTAC	GTATTTAAGA	GAATCTTGTC	ATGATAATTG	GGTACGACAT
		TATTCTAATG	CATAAATTCT	CTTAGAACAG	TACTATTAAAC	CCATGCTGTA
55	4451	AGTGATAAAT	GCTATTTGCG	ATCGTTACAT	AAAGTCAGTT	GGAAAGATGG
		TCACTATTTA	CGATAAAGCG	TAGCAATGTA	TTTCAGTCAA	CCTTTCFACC
	4501	ATTTGACAGA	TGTAACCTAA	TAGGTGCAAA	AATGTTAAAT	AACAGCATTC
		TAAACTGTCT	ACATTGAATT	ATCCACGTTT	TTACAATTTA	TTGTCGTAAG
	4551	TATCGGAAGA	TAGGATACCA	GTTATATTAT	ACAAAATCA	CTGGTTGGAT

		ATAGCCTTCT	ATCCTATGGT	CAATATAATA	TGTTTTTAGT	GACCAACCTA
4601		AAAACAGATT	CTGCAATATT	CGTAAAAGAT	GAAAGATTACT	GCGAATTTGT
		TTTTGTCTAA	GACGTTATAA	GCATTTTCTA	CTTCTAATGA	CGCTTAAACA
4651		AAACTATGAC	AATAAAAAGC	CATTTATCTC	AACGACATCG	TGTAATTCCT
5		TTTGATACTG	TTATTTTTCG	GTAAATAGAG	TTGCTGTAGC	ACATTAAGAA
		CCATGTTTTA	TGTATGTGTT	TCAGATATTA	TGAGATTACT	ATAAACTTTT
4701		GGTACAAAAT	ACATACACAA	AGTCTATAAT	ACTCTAATGA	TATTTGAAAA
		TGTATACTTA	TATTCGTAA	ACTATATTAA	TCATGAAGAA	AATGAAAAAG
4751		ACATATGAAT	ATAAGGCATT	TGATATAATT	AGTACTTCTT	TTACTTTTTTC
10		TATAGAAGCT	GTTTCACGAGC	GGTTGTTGAA	AACAACAAAA	TTATACATTC
		ATATCTTCGA	CAAGTGCTCG	CCAACAACCT	TTGTTGTTTT	AATATGTAAG
4801		AAGATGGCTT	ACATATACGT	CTGTGAGGCT	ATCATGGATA	ATGACAATGC
4851		TTCTACCGAA	TGTATATGCA	GACACTCCGA	TAGTACCTAT	TACTGTTACG
		ATCTCTAAAT	AGGTTTTTGG	ACAATGGATT	CGACCCTAAC	ACGGAATATG
4901		TAGAGATTTA	TCCAAAAACC	TGTTACCTAA	GCTGGGATTG	TGCCTTATAC
15		GTACTCTACA	ATCTCCTCTT	GAAATGGCTG	TAATGTTCAA	GAATACCGAG
		CATGAGATGT	TAGAGGAGAA	CTTTACCGAC	ATTACAAGTT	CTTATGGCTC
5001		GCTATAAAAA	TCTTGATGAG	GTATGGAGCT	AAACCTGTAG	TTACTGAATG
		CGATATTTTT	AGAACTACTC	CATACCTCGA	TTTGACATC	AATGACTTAC
20		CACAACTTCT	TGTCTGCATG	ATGCGGTGTT	GAGAGACGAC	TACAAAATAG
		GTGTTGAAGA	ACAGACGTAC	TACGCCACAA	CTCTCTGCTG	ATGTTTTATC
5101		TGAAAGATCT	GTTGAAGAAT	AACTATGTAA	ACAATGTTCT	TTACAGCGGA
		ACTTTCTAGA	CAACTTCTTA	TTGATACATT	TGTTACAAGA	AATGTCGCCT
5151		GGCTTTACTC	CTTTGTGTTT	GGCAGCTTAC	CTTAACAAAG	TTAATTTGGT
25		CCGAAATGAG	GAAACACAAA	CCGTGCAATG	GAATTGTTTC	AATTTAAACCA
		TAAACTTCTA	TTGGCTCATT	CGGCGGATGT	AGATATTTCA	AACACGGATC
5201		ATTTGAAGAT	AACCGAGTAA	GCCGCCTACA	TCATAAAAGT	TTGTGCCTAG
		GGTTAACTCC	TCTACATATA	GCCGTATCAA	ATAAAAATTT	AACAATGGTT
5251		CCAAATTGAGG	AGATGTATAT	CGGCATAGTT	TATTTTAAAA	TTGTTACCAA
30		AAACTTCTAT	TGAACAAAGG	TGCTGATACT	GACTTGCTGG	ATAACATGGG
		TTTGAAGATA	ACTTGTTTCC	ACGACTATGA	CTGAACGACC	TATTGTACCC
5301		ATGTATCTCT	TTAATGATCG	CTGTACAATC	TGGAAATATT	GAAATATGTA
5351		TACATGAGGA	AATTACTAGC	GACATGTTAG	ACCTTTATAA	CTTTTATACAT
		GCACACTACT	TAAAAAAAAT	AAAATGTCCA	GAACTGGGAA	AAATTGATCT
5401		CGTGTGATGA	ATTTTTTTTA	TTTACAGGT	CTTGACCCTT	TTTAACTAGA
35		TGCCAGCTGT	AATTCATGGT	AGAAAAGAAG	TGCTCAGGCT	ACTTTTCAAC
		ACGGTCGACA	TTAAGTACCA	TCTTTTCTTC	ACGAGTCCGA	TGAAAAGTTG
5501		AAAGGAGCAG	ATGTAAACTA	CATCTTTGAA	AGAAATGGAA	AATCATATAC
		TTTCCTCGTC	TACATTTGAT	GTAAGAACTT	TCTTTACCTT	TTAGTATATG
40		TGTTTTGGAA	TTGATTAAAG	AAAGTTACTC	TGAGACACAA	AAGAGGTAGC
		ACAAAACCTT	AACTAATTTT	TTTCAATGAG	ACTCTGTGTT	TTCTCCATCG
5601		TGAAGTGGTA	CTCTCAAAGG	TACGTGACTA	ATTAGCTATA	AAAAGGATCC
		ACTTCACCAT	GAGAGTTTCC	ATGCACTGAT	TAATCGATAT	TTTTCTTAGG
5651		TAGAGGATCA	TTATTTAACG	TAAACTAAAT	GGAAAAGCTA	TTTACAGGTA
45		ATCTCCTAGT	AATAAATTGC	ATTTGATTTA	CCTTTTCGAT	AAATGTCCAT
		CATACGGTGT	TTTCTGGAAT	CAAATGATTG	TGATTTTGAG	GATTTTATCA
5701		GTATGCCACA	AAAGACCTTA	GTTTACTAAG	ACTAAAACCTC	CTAAAATAGT
		ATACAATAAT	GACAGTGCTA	ACTGGTAAAA	AAGAAAAGCAA	ACAATTATCA
5751		TATGTTATTA	CTGTACGAT	TGACCATTTT	TTCTTTTCGTT	TGTTAATAGT
50		TGGCTAACAA	TTTTTATTAT	ATTTGTAGTA	TGCATAGTGG	TCTTTACGTT
		ACCGATTGTT	AAAAATAATA	TAAACATCAT	ACGTATCACC	AGAAAATGCAA
5801		TCTTTATTTA	AAGTTAATGT	GTTAAGATTA	AATGGAGTAA	TTGGATCCCC
5851		AGAAATAAAT	TTCAATTACA	CAATTCTAAT	TTACCTCATT	AACCTAGGGG
5901		CATCGATGGG	GAATTCACCTG	GCCGTGCTTT	TACAACGTCG	TGACTGGGAA
55		GTAGCTACCC	CTTAAGTGAC	CGGCAGCAAA	ATGTTGCAGC	ACTGACCCTT
		AACCCTGGCG	TTACCCAACCT	TAATCGCCTT	GCAGCACATC	CCCCTTTCGC
5951		TTGGGACCGC	AATGGGTTGA	ATTAGCGGAA	CGTCGTGTAG	GGGGAAGCG
6001		CAGCTGGCGT	AATAGCGAAG	AGGCCCGCAC	CGATCGCCCT	TCCCAACAGT
		GTCGACCGCA	TTATCGCTTC	TCCGGGCGTG	GCTAGCGGGA	AGGGTTGTCA

6051	TGCGCAGCCT	GAATGGCGAA	TGGCGCTTTG	CCTGGTTTCC	GGCACCAGAA
	ACGCGTCCGA	CTTACCGCTT	ACCGCGAAAC	GGACCAAAGG	CCGTGGTCTT
6101	GCGGTGCCGG	AAAGCTGGCT	GGAGTGCAT	CTTCCTGAGG	CCGATACTGT
	CGCCACGGCC	TTTCGACCGA	CCTCACGCTA	GAAGGACTCC	GGCTATGACA
5	6151	CGTCGTCCCC	TCAAACCTGGC	AGATGCACGG	TTACGATGCG
		GCAGCAGGGG	AGTTTGACCG	TCTACGTGCC	AATGCTACGC
	6201	CCAACGTAAAC	CTATCCCATT	ACGGTCAATC	CGCGGTTTGT
		GGTTGCATTG	GATAGGGTAA	TGCCAGTTAG	GCGGCAAACA
	6251	AATCCGACGG	GTTGTTACTC	GCTCACATTT	AATGTTGATG
10		TTAGGCTGCC	CAACAATGAG	CGAGTGTAAC	TTACAACCTAC
	6301	ACAGGAAGGC	CAGACGCGAA	TTATTTTGA	TGGCGTTAAC
		TGTCCTTCCG	GTCTGCGCTT	AATAAAACT	ACCGCAATTG
	6351	ATCTGTGGTG	CAACGGGCGC	TGGGTGCGTT	ACGGCCAGGA
		TAGACACCAC	GTTGCCCGCG	ACCCAGCCAA	TGCCGGTCCT
15	6401	CCGTCTGAAT	TTGACCTGAG	CGCATTTTFA	CGCGCCGGAG
		GGCAGACTTA	AACTGGACTC	CGGTAAAT	GCGCGGCTC
	6451	GCGGGTGATG	GTGCTGCGTT	GGAGTGACGG	CAGTTATCTG
		GCGCCACTAC	CACGACGCAA	CCTCACTGCC	GTCAATAGAC
	6501	ATATGTGGCG	GATGAGCGGC	ATTTTCCGTG	ACGTCTCGTT
20		TATACACCGC	CTACTCGCCG	TAAAGGCGAC	TGCAGAGCAA
	6551	CCGACTACAC	AAATCAGCGA	TTTCCATGTT	GCCACTCGCT
		GGCTGATGTG	TTTAGTCTGT	AAAGGTACAA	CGGTGAGCGA
	6601	TTTCAGCCGC	GCTGTACTGG	AGGCTGAAGT	TCAGATGTGC
		AAAGTCGGCG	CGACATGACC	TCCGACTTCA	AGTCTACACG
25	6651	GTGACTACCT	ACGGGTAAAC	GTTTCTTTAT	GGCAGGGTGA
		CACTGATGGA	TGCCCATTGT	CAAAGAAATA	CCGTCCCACT
	6701	GCCAGCGGCA	CCGCGCTTTT	CGGCGGTGAA	ATTATCGATG
		CGGTGCGCGT	GGCGCGGAAA	GCCGCCACTT	TAATAGCTAC
	6751	TTATGCCGAT	CGCGTCACAC	TACGTCTGAA	CGTCGAAAAC
30		AATACGGCTA	GCGCAGTGTG	ATGCAGACTT	GCAGCTTTTG
	6801	GGAGCGCCGA	AATCCCGAAT	CTCTATCGTG	CGGTGGTTGA
		CCTGCGGGCT	TTAGGGCTTA	GAGATAGCAC	GCCACCAACT
	6851	GCCGACGGCA	CGCTGATTGA	AGCAGAAGCC	TGCGATGTGC
		CGGTGCGCGT	GCGACTAACT	TGCTCTTCGG	ACGCTACAGC
35	6901	GGTGCGGATT	GAAAAATGGT	TGCTGCTGCT	GAACGGCAAG
		CCACGCCTAA	CTTTTACCAG	ACGACGACGA	CTTGCCGTTT
	6951	TTTCGAGCGT	TAACCGTCAC	GAGCATCATC	CTCTGATGCG
		AAGTCCGCA	ATTGGCAGTG	CTCGTAGTAG	GAGACGTACC
	7001	GATGAGCAGA	CGATGGTGCA	GGATATCCTG	CTGATGAAGC
40		CTACTCGTCT	GCTACCACGT	CCTATAGGAC	GACTACTTCG
	7051	TAACGCCGTG	CGCTGTTGCG	ATTATCCGAA	CCATCCGCTG
		ATTGCGGCAC	GCGACAAGCG	TAATAGGCTT	GGTAGGCGAC
	7101	TGTGCGACCG	CTACGGCCTG	TATGTGGTGG	ATGAAGCCAA
		ACACGCTGGC	GATGCCGGAC	ATACACCACC	TACTTCGGTT
45	7151	CACGGCATGG	TGCCAATGAA	TGCTCTGACC	GATGATCCGC
		GTGCCGTACC	ACGGTTACTT	AGCAGACTGG	CTACTAGGCG
	7201	GGCGATGAGC	GAACGCGTAA	CGCGAATGGT	GCAGCGCGAT
		CCGCTACTCG	CTTGCGCATT	GCGCTTACCA	CGTCGCGCTA
	7251	CGAGTGTGAT	CATCTGGTCT	CTGGGGAATG	AATCAGGCCA
50		GCTCACACTA	GTAGACCAGC	GACCCCTTAC	TTAGTCCGGT
	7301	CACGACGCGC	TGTATCGCTG	GATCAAATCT	GTGATCCCTT
		GTGCTGCGCG	ACATAGCGAC	CTAGTTTGA	CAGCTAGGAA
	7351	GCAGTATGAA	GGCGGCGGAG	CCGACACCAC	GGCCACCGAT
		CGTCATACTT	CCGCCGCCTC	GGCTGTGGTG	CCGTGGGCTA
55	7401	CGATGTACGC	GCGCGTGGAT	GAGACACAGC	CCTTCCCGGC
		GCTACATGCG	CGCGCACCTA	CTTCTGGTCT	GGAAGGGCCG
	7451	TGTTCCATCA	AAAAATGGCT	TTCGCTACCT	GGAGAGACGC
		ACCAGGTAGT	TTTTTACCGA	AAGCGATGGA	CCTCTCTGCG
	7501	CCTTTGCAGAA	TACGCCACAG	CGATGGGTAA	CAGTCTTGGC

		GGAAACGCTT	ATGCGGGTGC	GCTACCCATT	GTCAGAACCG	CCAAAGCGAT
7551		AATACTGGCA	GGCGTTTCGT	CAGTATCCCC	GI'TTACAGGG	CGGCTTCGTC
		TTATGACCGT	CCGCAAAGCA	GTCATAGGGG	CAAATGTCCC	GCCGAAGCAG
7601		TGGGACTGGG	TGGATCAGTC	GCTGATTAAA	TATGATGAAA	ACGGCAACCC
5		ACCCTGACCC	ACCTAGTCAG	CGACTAATTT	ATACTACTTT	TGCCGTTGGG
7651		GTGGTCGGCT	TACGGCGGTG	ATTTTGGCGA	TACGCCGAAC	GATCGCCAGT
		CACCAGCCGA	ATGCCGCCAC	TAAAACCGCT	ATGCGGCTTG	CTAGCGGTCA
7701		TCTGTATGAA	CGGTCTGGTC	TTTGCCGACC	GCACGCCGCA	TCCAGCGCTG
		AGACATACTT	GCCAGACCAG	AAACGGCTGG	CGTGCGGCGT	AGGTCGCGAC
10	7751	ACGGAAGCAA	AACACCAGCA	GCAGTTTTTC	CAGTTCCGTT	TATCCGGGCA
		TGCCTTTCGTT	TTGTGGTCGT	CGTCAAAAAG	GTCAAGGCAA	ATAGGCCCGT
7801		AACCATCGAA	GTGACCAGCG	AA'TACCTGTT	CCGTATAGC	GATAACGAGC
		TTGGTAGCTT	CACTGGTCGC	TTATGGACAA	GGCAGTATCG	CTATTGCTCG
7851		TCCTGCACTG	GATGGTGGCG	CTGGATGGTA	AGCCGCTGGC	AAGCGGTGAA
15		AGGACGTGAC	CTACCACCGC	GACCTACCAT	TCGGCGACCG	TTCGCCACTT
7901		GTGCCCTCTGG	ATGTCGCTCC	ACAAGGTAAA	CAGTTGATTG	AACTGCCTGA
		CACGGAGACC	TACAGCGAGG	TGTTCCATTT	GTCAACTAAC	TTGACGGACT
7951		ACTACCGCAG	CCGGAGAGCG	CCGGGCAACT	CTGGCTCACA	GTACGCGTAG
		TGATGGCGTC	GGCCTCTCGC	GGCCCCGTGA	GACCGAGTGT	CATGCGCATC
20	8001	TGCAACCGAA	CGCGACCGCA	TGTCAGAAAG	CCGGGCACAT	CAGCGCCTGG
		ACGTTGGCTT	GCGCTGGCGT	ACCAGTCTTC	GGCCCCGTGA	GTGCGGGACC
8051		CAGCATGGGC	GTCTGGCGGA	AAACCTCAGT	GTGACGCTCC	CCGCCGCGTC
		GTCGTACCGG	CAGACCGCCT	TTTGGAGTCA	CACTGCGAGG	GGCGGCGCAG
8101		CCACGCCATC	CCGCATCTGA	CCACCAGCGA	AATGGATTTT	TGCATCGAGC
25		GGTGCGGTAG	GGCGTAGACT	GGTGGTCGCT	TTACCTAAAA	ACGTAGCTCG
8151		TGGGTAATAA	GCGTTGGCAA	TTTAACCGCC	AGTCAGGCTT	TCTTTCACAG
		ACCCATTATT	CGCAACCGTT	AAATTGGCGG	TCAGTCCGAA	AGAAAGTGTC
8201		ATGTGGATTG	GCGATAAAAA	ACAACCTGCTG	ACGCCGCTGC	GCGATCAGTT
		TACACCTAAC	CGCTATTTTT	TGTTGACGAC	TGCGGCGACG	CGCTAGTCAA
30	8251	CACCCGTGCA	CCGCTGGATA	ACGACATTGG	CGTAAGTGAA	GCGACCCGCA
		TGGGGCACGT	GGCGACCTAT	TGCTGTAACC	GCATTCACTT	CGCTGGGCGT
8301		TTGACCCCTAA	CGCCTGGGTC	GAACGCTGGA	AGGCGCGGGG	CCATTACCAG
		AACTGGGATT	GCGGACCCAG	CTTGCGACCT	TCCGCCGCCC	GGTAATGGTC
8351		GCCGAAGCAG	CGTTGTTGCA	GTGCACGGCA	GATACACTTG	CTGATGCGGT
35		CGGCTTCGTC	GCAACAACGT	CACGTGCCGT	CTATGTGAAC	GACTACGCCA
8401		GCTGATTACG	ACCGCTCACG	CGTGGCAGCA	TCAGGGGAAA	ACCTTATTTA
		CGACTAATGC	TGGCGAGTGC	GCACCGTCGT	AGTCCCCCTT	TGGAATAAAT
8451		TCAGCCGGAA	AACCTACCGG	ATTGATGGTA	GTGGTCAAAT	GGCGATTACC
		AGTCGGCCTT	TTGGATGGCC	TAAC'TACCAT	CACCACTTTA	CCGCTAATGG
40	8501	GTTGATGTTG	AAGTGGCGAG	CGATACACCG	CATCCGGCGC	GGATTGGCCT
		CAACTACAAC	TTCACCGCTC	GCTATGTGGC	GTAGGCCGCG	CCTAACCGGA
8551		GAAC'TGCCAG	CTGGCGCAGG	TAGCAGAGCG	GGTAAACTGG	CTCGGATTAG
		CTTGACGGTC	GACCGCGTCC	ATCGTCTCGC	CCATTTGACC	GAGCCTAATC
8601		GGCCGCAAGA	AAACTATCCC	GACCGCCTTA	CTGCCGCTTG	TTTTGACCGC
45		CCGGCGTICT	TTTGATAGGG	CTGGCGGAAT	GACGGCGGAC	AAAAC'TGGCG
8651		TGGGATCTGC	CATTGTCTAGA	CATGTATACC	CCGTACGCTC	TCCCGAGCGA
		ACCCTAGACG	GTAACAGTCT	GTACATATGG	GGCATGCAGA	AGGGCTCGCT
8701		AAACGGTCTG	CGCTGCGGGA	CGCGCGAATT	GAATTATGGC	CCACACCAGT
		TTTGCCAGAC	GCGACGCCCT	GCGCGCTTAA	CTTAATACCG	GGTGTGGTCA
50	8751	GGCGCGCGCA	CTTCCAGTTC	AACATCAGCC	GCTACAGTCA	ACAGCAACTG
		CCGCGCCGCT	GAAGGTCAAG	TTGTAGTCGG	CGATGTCAGT	TGTCGTTGAC
8801		ATGGAAACCA	GCCATCGCCA	TCTGCTGCAC	GCGGAAGAAG	GCACATGGCT
		TACCTTTGGT	CGGTAGCGGT	AGACGACGTG	CGCCTTCTTC	CGTGTACCGA
8851		GAATATCGAC	GGTTTCCATA	TGGGGATTGG	TGGCGACGAC	TCCTGGAGCC
55		CTTATAGCTG	CCAAAGGTAT	ACCCCTAACC	ACCGCTGCTG	AGGACCTCGG
8901		CGTCAGTATC	GGCGGAATTC	CAGCTGAGCG	CCGGTCGCTA	CCATTACCAG
		GCAGTCATAG	CCGCCTTAAG	GTCGACTCGC	GGCCAGCGAT	GGTAATGGTC
8951		TTGGTCTGGT	GTCAAAAATA	ATAATAACCG	GGCAGGGGGG	ATCCGGAGCT
		AACCAGACCA	CAGTTTTTAT	TATTATTGGC	CCGTCCCCC	TAGGCCTCGA

9001	TATCGCAGAT	CAATGATCGC	TGTACAATCT	GGAAATATTG	AAATATGTAG
	ATAGCGTCTA	GTTACTAGCG	ACATGTTAGA	CCTTTATAAC	TTTATACATC
9051	CACACTACTT	AAAAAAAATA	AAATGTCCAG	AACTGGGAAA	AATTGATCTT
	GTGTGATGAA	TTTTTTTTTAT	TTTACAGGTC	TTGACCCTTT	TTAAC TAGAA
5	9101	GCCAGCTGTA	ATTCATGGTA	GAAAAGAAGT	GCTCAGGCTA
		CGGTCGACAT	TAAGTACCAT	CTTTTCTTCA	CGAGTCCGAT
	9151	AAGGAGCAGA	TGTAAACTAC	ATCTTTGAAA	GAAATGGAAA
		TTCTCTCGTCT	ACATTTGATG	TAGAAACTTT	CTTTACCTTT
	9201	GTTTTGGAAT	TGATTAAAGA	AAGTTACTCT	GAGACACAAA
10		CAAAACCTTA	ACTAATTTCT	TTCAATGAGA	CTCTGTGTTT
	9251	GAAGTGGTAC	TCTCAAAGGT	ACGTGACTAA	TTAGCTATAA
		CTTCACCATG	AGAGTTTCCA	TGCACTGATT	AATCGATATT
	9301	GTACCCTCGA	GTCIAGAATC	GATCCCGGGT	TAATTAATTA
		CATGGGAGCT	CAGATCTTAG	CTAGGGCCCA	ATTAATTAAT
15	9351	AAGGTGAAAA	CGAAACTATT	TGTAGCTTAA	TAAATTAGAG
		TTCCACTTTT	GCTTTGATAA	ACATCGAATT	AATTAATCTC
	9401	CTATACTTAA	AAAGTGAAAA	TAAATACAAA	GGTTCTTGAG
		GATATGAATT	TTTCACTTTT	ATTTATGTTT	CCAAGAACTC
	9451	AATTGAAAGC	GAGAAATAAT	CATAAATTAT	TTCATTATCG
20		TTAACTTTTCG	CTCTTTATTA	GTATTTAATA	AAGTAATAGC
	9501	TAAGTTTGTA	TCGTA		GCTATAGGCA
		ATTCAAACAT	AGCAT		

FIGURE 6